Packing and Embedding Large subgraphs

By

Padraig Condon

A thesis submitted to
the University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Mathematics
College of Engineering and Physical Sciences
University of Birmingham
June 2020

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

ABSTRACT

This thesis contains several embedding results for graphs in both random and non random settings. Most notably, we resolve a long standing conjecture that the threshold probability for Hamiltonicity in the random binomial subgraph of the hypercube equals 1/2.

In Chapter 2 we obtain the following perturbation result regarding the hypercube $\mathcal{Q}^n$: if $H \subseteq \mathcal{Q}^n$ satisfies $\delta(H) \geq \alpha n$ with $\alpha > 0$ fixed and we consider a random binomial subgraph $\mathcal{Q}^n_p$ of $\mathcal{Q}^n$ with $p \in (0, 1]$ fixed, then with high probability $H \cup \mathcal{Q}^n_p$ contains k edge-disjoint Hamilton cycles, for any fixed $k \in \mathbb{N}$. This result is part of a larger volume of work where we also prove the corresponding hitting time result for Hamiltonicity.

In Chapter 3 we move to a non random setting. Rather than pack a small number of Hamilton cycles into a fixed host graph, our aim is to achieve optimally sized packings of more general families of graphs. More specifically, we provide a degree condition on a regular n-vertex graph G which ensures the existence of a near optimal packing of any family $\mathcal{H}$ of bounded degree n-vertex k-chromatic separable graphs into G. In particular, this yields approximate versions of the the tree packing conjecture, the Oberwolfach problem, the Alspach problem and the existence of resolvable designs in the setting of regular host graphs of high degree.

ACKNOWLEDGEMENTS

To my supervisors, Deryk Osthus and Daniela Kühn, thank you for your time, support, patience and guidance. Thank you for introducing me to a wonderful world that I would never otherwise have experienced.

To my coauthors, Jaehoon, António and Alberto, thank you for all of your help, time and engaging discussions. Thank you for all of the laughs.

To my friends in the department, thank you for your help and support, for the fun lunch time discussions and coffee breaks and for all of the memories.

To my maths friends from Trinity, Keith, Aran, Nastia, Masha and Tim, thank you for helping me to fall in love with mathematics.

To my parents John and Oonagh, and to Eoin, Orlaith and Orlagh, thank you for your love and support always.

PUBLICATIONS AND SUBMISSIONS

During my PhD I have written the following four papers, two of which are the subject of this thesis.

P. Condon, A. Espuny Díaz, A. Girão, D. Kühn and D. Osthus, **Hamiltonicity of random subgraphs of the hypercube**, submitted for publication.
The ideas in this paper were contributed to equally by all authors. The first author was responsible for writing the section 'Near-spanning trees in random subgraphs of the hypercube', which is included in this thesis. The second author was responsible for writing the section 'Tiling random subgraphs of the hypercube with small cubes.' This section is not included in this thesis due to word constraints. The writing of the remaining sections was shared equally by the first and second authors. The section 'Hitting time result' is also left out due to word constraints. The remaining sections make up Chapter 2 of this thesis.

P. Condon, J. Kim, D. Kühn and D. Osthus, **A bandwidth theorem for approximate decompositions**, Proceedings London Mathematical Society 118 (2019), 1393-1449.
The ideas in this paper were contributed to equally by all authors. The writing was shared equally between the first and second authors. This paper makes up Chapter 3 of this thesis.

P. Condon, A. Espuny Díaz, J. Kim, D. Kühn, and D. Osthus, **Resilient degree sequences with respect to Hamilton cycles and matchings in random graphs**, Electronic Journal of Combinatorics 26 (2019), P4.54.
The ideas in this paper were contributed to equally by all authors. The writing was shared equally between the first and second authors. Due to word constraints this paper has not

been included in this thesis, however the main results are described Chapter 1.

P. Condon, A. Espuny Díaz, A. Girão, D. Kühn, and D. Osthus, **Dirac's theorem for random regular graphs**, to appear in Combinatorics, Probability and Computing.
The ideas in this paper were contributed to equally by all authors. The writing was shared equally between the first and second authors. Once more, due to word constraints this paper is not included in this thesis, however the main results are described Chapter 1.

Contents

Chapter 1

Introduction

The topics of this thesis are extremal graph theory and probabilistic combinatorics. Our results all concern the following question: given two graphs G and H as input, determine whether G contains a subgraph isomorphic to H?

We are interested in the case where this subgraph is spanning in G. Our goal will be to determine the existence of such subgraphs and not to find specific examples. In general, even this is an NP-complete problem. One famous example which is known to be NP-complete is the problem of determining whether G contains a cycle that covers every vertex exactly once (i.e. a Hamilton cycle). As a result of this NP-completeness, for many classes of graphs G and H, the study of this question has moved in the direction of finding sufficient conditions, particularly in the form of degree conditions. The classic example concerning Hamilton cycles is the theorem of Dirac, which states that for $n \geq 3$, every n-vertex graph with minimum degree at least $n/2$ is Hamiltonian. We study this question of containing a given subgraph for the case of Hamilton cycles, but also for more general families of graphs. Moreover, we tackle this problem in both random and non random settings.

We begin in the random setting. In graph theory, a random graph is a graph that has been sampled via some probability distribution over a fixed collection of graphs. The most well-known and studied random graph model is the binomial model $G_{n,p}$. One can generate a random graph $G \sim G_{n,p}$ according to this distribution by starting with the n-vertex complete graph and deleting each edge with probability $1-p$, independently of every other edge. Given a property of interest, the aim of study is to determine the threshold probability p^* for this model beyond which the property is realised with high probability (sometimes referred to as the phase transition).

The threshold probability for Hamilton cycles appearing in the binomial model has been very well understood since the 1970s. Another random graph model which can be sampled in a similar way to the binomial model, is the model of random subgraphs of the hypercube. This time, instead of starting with the complete graph, one starts with the n-dimensional hypercube. In contrast to the binomial model, the problem of determining the threshold probability for Hamiltonicity in random subgraphs of the hypercube has proven much more difficult. In Chapter 2 we resolve this question by proving the threshold probability occurs at $p^* = 1/2$. Moreover, instead of embedding just a single Hamilton cycle at this threshold, we actually prove that for any fixed $k \in \mathbb{N}$, the same threshold holds for packing k Hamilton cycles into our random subgraph of the hypercube. Here we say that k Hamilton cycles pack into a graph G whenever G contains k edge-disjoint Hamilton cycles as a subgraph.

In Chapter 3 we switch to a non random setting. Instead of packing a small number of Hamilton cycles into a fixed host graph G, our aim is now to pack the optimal number. In fact for the case of Hamilton cycles this problem was already well understood: recently Csaba, Kühn, Lo, Osthus, and Treglown [35] showed that any r-regular n-vertex graph G with $r \geq \lfloor n/2 \rfloor$ has a decomposition into Hamilton cycles and at most 1 perfect matching. Here we say that a graph G decomposes into a collection of graphs $\mathcal{H}$ in the special case where $\mathcal{H}$ packs into G and covers every edge of G (i.e. $\mathcal{H}$ and G have the same number of edges). In Chapter 3, we allow for a more general class of graphs than Hamilton cycles, characterised by the graph property of having small bandwidth. Here a graph has small bandwidth if each vertex only has edges to vertices that are 'nearby' with respect to some vertex ordering (this includes Hamilton cycles). We provide a degree condition on a regular graph G which is sufficient to ensure the existence of a near optimal packing of such families of graphs into G. In general, this degree condition is best possible.

Our result in Chapter 3 can be viewed as a near optimal packing version of the bandwidth theorem of Böttcher, Schacht and Taraz [24] (which concerns embedding a single small bandwidth graph) in the setting of regular host graphs. Several design-type corollaries follow. One example concerns the tree packing conjecture of Gyárfás and Lehel, which asks for decompositions of the complete graph into specific collections of trees. Our result yields an approximate version of this conjecture for bounded degree trees, in the setting of regular host graphs of high degree. Another long standing conjecture, the Oberwolfach problem, which concerns decomposing the complete graph into cycles, was recently solved by Glock, Joos, Kim, Kühn, and Osthus [48], where our result was an important tool in their proof.

Finally, we mention some further work that has not been included in this thesis due to word constraints. The local resilience of a graph G with respect to a property P measures how much one has to change G locally in order to destroy P. In a sequence of two papers [33, 32] we prove 'resilience' versions of several classical results for the binomial and random regular graph models. Here we sample an n-vertex d-regular graph according to the random regular model $G_{n,d}$ by choosing a graph uniformly at random from the collection of all n-vertex d-regular graphs.

In [32] we solve a conjecture of Ben-Shimon, Krivelevich and Sudakov by proving a resilience version of Dirac's theorem in the setting of random regular graphs. More precisely, we show that, whenever d is sufficiently large compared to $\varepsilon > 0$, with high probability the following holds: let G' be any subgraph of the random n-vertex d-regular graph $G_{n,d}$ with minimum degree at least $(1/2 + \varepsilon)d$. Then G' is Hamiltonian. Here the condition that d is large is necessary.

In [33] we consider strengthenings of Dirac's theorem. Pósa's theorem states that any n-vertex graph G whose degree sequence $d_1 \leq \ldots \leq d_n$ satisfies $d_i \geq i + 1$ for all $i < n/2$ has a Hamilton cycle. This result is generalised further by Chvátal's theorem, which characterises

all degree sequences that ensure the existence of a Hamilton cycle. We prove a resilience version of Pósa's Hamiltonicity condition for the binomial model $G_{n,p}$ and show that a natural guess for a resilient version of Chvátal's theorem for this model fails to be true.

Chapter 2

Hamiltonicity of random subgraphs of the hypercube

2.1 Introduction

The n-dimensional *hypercube* $\mathcal{Q}^n$ is the graph whose vertex set consists of all n-bit 01-strings, where two vertices are joined by an edge whenever their corresponding strings differ by a single bit. The hypercube and its subgraphs have attracted much attention in graph theory and computer science, e.g. as a sparse network model with strong connectivity properties. It is well known that hypercubes contain spanning paths (also called *Gray codes* or *Hamilton paths*) and, for all $n \geq 2$, they contain spanning cycles (also referred to as *cyclic Gray codes* or *Hamilton cycles*). Classical applications of Gray codes in computer science are described in the surveys of Savage [89] and Knuth [65]. Applications of hypercubes to parallel computing are discussed in the monograph of Leighton [82].

2.1.1 Spanning subgraphs in hypercubes

The systematic study of spanning paths, trees and cycles in hypercubes was initiated in the 1970's. There is by now an extensive literature about subtrees of the hypercube; see, for instance, results of Bhatt, Chung, Leighton, and Rosenberg [10] about embedding subdivided trees (instigated by processor allocation in distributed computing systems).

As a generalization of Hamilton paths, Caha and Koubek [27] considered the problem of finding a collection of spanning vertex-disjoint paths, given a prescribed set of endpoints. After several improvements [29, 52], this problem was recently resolved by Dvořák, Gregor, and Koubek [40].

The applications of hypercubes as networks in computer science inspired questions about the reliability of its properties. This led to considering 'faulty' hypercubes in which some edges or vertices are missing. For instance, Chan and Lee [28] showed that, if $\mathcal{Q}^n$ has at most $2n-5$ faulty edges and every vertex has (non-faulty) degree at least 2, then there is a Hamilton cycle in $\mathcal{Q}^n$ which avoids all faulty edges (and this condition is best possible). They also showed that the general problem of determining the Hamiltonicity of $\mathcal{Q}^n$ with a larger number of faulty edges is NP-complete. More generally, Dvořák and Gregor [39] studied the existence of spanning collections of vertex-disjoint paths with prescribed endpoints in faulty hypercubes. (We will apply these results in our proofs, see Section 2.7.3 for details.) These can be seen as extremal results about the *robustness* of the hypercube with respect to containing spanning collections of paths and cycles.

2.1.2 Hamilton cycles in binomial random graphs

One of the most studied random graph models is the binomial random graph $G_{n,p}$. Here we have a (labelled) set of n vertices and we include each edge with probability p independently of all other edges.

Given some monotone increasing graph property $\mathcal{P}$, a function $p^* = p^*(n)$ is said to be a (coarse) *threshold* for $\mathcal{P}$ if $\mathbb{P}[G_{n,p} \in \mathcal{P}] \to 1$ whenever $p/p^* \to \infty$ and $\mathbb{P}[G_{n,p} \in \mathcal{P}] \to 0$ whenever $p/p^* \to 0$. One can define the stronger notion of a sharp threshold similarly: $p^* = p^*(n)$ is said to be a *sharp threshold* for $\mathcal{P}$ if, for all $\varepsilon > 0$, we have that $\mathbb{P}[G_{n,p} \in \mathcal{P}] \to 1$

whenever $p \geq (1+\varepsilon)p^*$ and $\mathbb{P}[G_{n,p} \in \mathcal{P}] \to 0$ whenever $p \leq (1-\varepsilon)p^*$. The problem of finding the threshold for the containment of a Hamilton cycle was solved independently by Pósa [87] and Koršunov [73]. Furthermore, Koršunov [73] determined the sharp threshold for Hamiltonicity to be $p^* = \log n/n$. These results were later made even more precise by Komlós and Szemerédi [72]. It is worth noting that $p^* = \log n/n$ is also the sharp threshold for the property of having minimum degree at least 2. In this sense, the results about Hamilton cycles in $G_{n,p}$ can be interpreted as saying that the natural obstruction of having sufficiently high minimum degree is also an 'almost sufficient' condition.

A property that generalises Hamiltonicity is that of containing k edge-disjoint Hamilton cycles, for some $k \in \mathbb{N}$. We will present more results in this direction in Section 2.1.4; for now, let us simply note that the sharp threshold for the containment of k edge-disjoint Hamilton cycles in $G_{n,p}$, for some $k \in \mathbb{N}$ independent of n, is $p^* = \log n/n$, i.e. the same as the threshold for Hamiltonicity.

The study of *robustness* of graph properties has also attracted much attention recently. For instance, given a graph G which is known to satisfy some property $\mathcal{P}$, consider a random subgraph G_p obtained by deleting each edge of G with probability $1-p$, independently of all other edges. The problem then is to determine the range of p for which G_p satisfies $\mathcal{P}$ with high probability. In this setting, a result of Krivelevich, Lee, and Sudakov [75] asserts that, for any n-vertex graph G with minimum degree at least $n/2$, the graph G_p is asymptotically almost surely Hamiltonian whenever $p \gg \log n/n$. This can be viewed as a robust version of Dirac's theorem on Hamilton cycles.

2.1.3 Hamilton cycles in binomial random subgraphs of the hypercube

Throughout this paper, we will consider random subgraphs of the hypercube and show that the hypercube is robustly Hamiltonian in the above sense. We will denote by $\mathcal{Q}^n_p$ the random subgraph of the hypercube obtained by removing each edge of $\mathcal{Q}^n$ with probability $1-p$ independently of every other edge.

The random graph $\mathcal{Q}^n_p$ was first studied by Burtin [26], who proved that the sharp threshold for connectivity is $1/2$. This result was later made more precise by Erdős and Spencer [42] and Bollobás [13]. As a related problem, Dyer, Frieze, and Foulds [41] determined the sharp threshold for connectivity in subgraphs of $\mathcal{Q}^n$ obtained by removing both vertices and edges uniformly at random. Later, Bollobás [15] proved that $1/2$ is also the sharp threshold for the containment of a perfect matching in $\mathcal{Q}^n_p$. As with the $G_{n,p}$ model, this also coincides with the threshold for having minimum degree at least 1.

The main goal of this paper is to study the analogous problem for Hamiltonicity in random subgraphs of the hypercube. There is a folklore conjecture that the sharp threshold for Hamiltonicity in $\mathcal{Q}^n_p$ should be $1/2$, i.e. the same as the threshold for having minimum degree at least 2. This question was explicitly asked by Bollobás [16] at several conferences in the 1980's, in the ICM surveys of Frieze [46] and Kühn and Osthus [80], as well as the recent survey of Frieze [47]. A special case of our first result resolves this problem.

Theorem 2.1.1. *For any $k \in \mathbb{N}$, the sharp threshold for the property of containing k edge-disjoint Hamilton cycles in $\mathcal{Q}^n_p$ is $p^* = 1/2$.*

For $k = 1$, this can be seen as a probabilistic version of the result on faulty hypercubes [28], and also as a statement about the robustness of Hamiltonicity in the hypercube.

While, for $p < 1/2$, with high probability $\mathcal{Q}^n_p$ will not contain a Hamilton cycle, it turns out that the reason for this is mostly due to local obstructions (e.g., vertices with degree zero or one). More precisely, we prove that, for any constant $p \in (0, 1/2)$, a.a.s. the random graph $\mathcal{Q}^n_p$ contains an almost spanning cycle.

Theorem 2.1.2. *For any $\delta, p \in (0, 1]$, a.a.s. the graph $\mathcal{Q}^n_p$ contains a cycle of length at least $(1-\delta)2^n$.*

We believe that the probability bound is far from optimal, in the sense that random subgraphs of the hypercube where edges are picked with vanishing probability should also satisfy this property.

Conjecture 2.1.3. *Suppose that* $p = p(n)$ *satisfies that* $pn \to \infty$*. Then, a.a.s.* $\mathcal{Q}^n_p$ *contains a cycle of length* $(1 - o(1))2^n$*.*

Similarly, it would be interesting to determine which (long) paths and (almost spanning) trees can be found in $\mathcal{Q}^n_p$. Moreover, our methods might also be useful to embed other large subgraphs, such as F-factors.

Conjecture 2.1.4. *Suppose* $\varepsilon > 0$ *and an integer* $\ell \geq 2$ *are fixed and* $p \geq 1/2 + \varepsilon$*. Then, a.a.s.* $\mathcal{Q}^n_p$ *contains a* C_{2^ℓ}*-factor, that is, a set of vertex-disjoint cycles of length* 2^ℓ *which together contain all vertices of* $\mathcal{Q}^n$*.*

2.1.4 Hitting time results

Remarkably, the above intuition that having the necessary minimum degree is an 'almost sufficient' condition for the containment of edge-disjoint perfect matchings and Hamilton cycles can be strengthened greatly via so-called hitting time results. These are expressed in terms of random graph processes. The general setting is as follows. Let G be an n-vertex graph with $m = m(n)$ edges, and consider an arbitrary labelling $E(G) = \{e_1, \ldots, e_m\}$. The *$G$-process* is defined as a random sequence of nested graphs $\tilde{G}(\sigma) = (G_t(\sigma))_{t=0}^m$, where σ is a permutation of $[m]$ chosen uniformly at random and, for each $i \in [m]_0$, we set $G_i(\sigma) = (V(G), E_i)$, where $E_i \coloneqq \{e_{\sigma(j)} : j \in [i]\}$. Given any monotone increasing graph property $\mathcal{P}$ such that $G \in \mathcal{P}$, the *hitting time* for $\mathcal{P}$ in the above G-process is the random variable $\tau_{\mathcal{P}}(\tilde{G}(\sigma)) \coloneqq \min\{t \in [m]_0 : G_t(\sigma) \in \mathcal{P}\}$.

Let us denote the properties of containing a perfect matching by $\mathcal{PM}$, Hamiltonicity by $\mathcal{HAM}$, and connectivity by $\mathcal{CON}$, respectively. For any $k \in \mathbb{N}$, let $\boldsymbol{\delta k}$ denote the property of having minimum degree at least k, and let $\mathcal{HM}k$ denote the property of containing $\lfloor k/2 \rfloor$ edge-disjoint Hamilton cycles and, if k is odd, one matching of size $\lfloor n/2 \rfloor$ which is edge-disjoint from these Hamilton cycles. With this notion of hitting times, many of the results about thresholds presented in Sections 2.1.2 and 2.1.3 can be strengthened significantly. For instance, Bollobás and Thomason [18] showed that a.a.s. $\tau_{\mathcal{CON}}(\tilde{K_n}(\sigma)) = \tau_{\boldsymbol{\delta}1}(\tilde{K_n}(\sigma))$ and, if n is even, then a.a.s. $\tau_{\mathcal{PM}}(\tilde{K_n}(\sigma)) = \tau_{\boldsymbol{\delta}1}(\tilde{K_n}(\sigma))$. Ajtai, Komlós, and Szemerédi [2] and Bollobás [14] independently proved that a.a.s. $\tau_{\mathcal{HAM}}(\tilde{K_n}(\sigma)) = \tau_{\boldsymbol{\delta}2}(\tilde{K_n}(\sigma))$. This was later generalised by Bollobás and Frieze [17], who proved that, given any $k \in \mathbb{N}$, for n even a.a.s. $\tau_{\mathcal{HM}k}(\tilde{K_n}(\sigma)) = \tau_{\boldsymbol{\delta}k}(\tilde{K_n}(\sigma))$.

A hitting time result for the property of having k edge-disjoint Hamilton cycles when k is allowed to grow with n is still not known, even in K_n-processes. As a slightly weaker notion, consider property $\mathcal{H}$, where we say that a graph G satisfies property $\mathcal{H}$ if it contains $\lfloor \delta(G)/2 \rfloor$ edge-disjoint Hamilton cycles, together with an additional edge-disjoint matching of size $\lfloor n/2 \rfloor$ if $\delta(G)$ is odd. Knox, Kühn, and Osthus [64], Krivelevich and Samotij [76] as well as Kühn and Osthus [81] proved results for different ranges of p which, together, show that $G_{n,p}$ a.a.s. satisfies property $\mathcal{H}$.

For graphs other than the complete graph, Johansson [59] recently obtained a robustness version of the hitting time results for Hamiltonicity. In particular, for any n-vertex graph G with $\delta(G) \geq (1/2+\varepsilon)n$, he proved that a.a.s. $\tau_{\mathcal{HAM}}(\tilde{G}(\sigma)) = \tau_{\boldsymbol{\delta}2}(\tilde{G}(\sigma))$. This was later extended to a larger class of graphs G and to hitting times for $\mathcal{HM}2k$, for all $k \in \mathbb{N}$ independent of n, by Alon and Krivelevich [6].

In the setting of random subgraphs of the hypercube, Bollobás [15] determined the hitting time for perfect matchings by showing that a.a.s. $\tau_{\mathcal{PM}}(\tilde{\mathcal{Q}^n}(\sigma)) = \tau_{\mathcal{CON}}(\tilde{\mathcal{Q}^n}(\sigma)) = \tau_{\boldsymbol{\delta}1}(\tilde{\mathcal{Q}^n}(\sigma))$. One of our main results (which implies Theorem 2.1.1) is a hitting time result for Hamiltonicity (and, more generally, property $\mathcal{HM}k$) in $\mathcal{Q}^n$-processes. Again, this question was raised by Bollobás [16] at several conferences.

Theorem 2.1.5. *For all $k \in \mathbb{N}$, a.a.s. $\tau_{\mathcal{HM}k}(\tilde{\mathcal{Q}^n}(\sigma)) = \tau_{\boldsymbol{\delta}k}(\tilde{\mathcal{Q}^n}(\sigma))$, that is, the hitting time for the containment of a collection of $\lfloor k/2 \rfloor$ Hamilton cycles and $k-2\lfloor k/2 \rfloor$ perfect matchings,* *all pairwise edge-disjoint, in $\mathcal{Q}^n$-processes is a.a.s. equal to the hitting time for the property of having minimum degree at least k.*

We omit the proof of Theorem 2.1.5 due to word constraints. A sketch is provided in Section 2.2.6 and the full proof can be found in [31]. We also wonder whether this is true if k is allowed to grow with n, and propose the following conjecture which, if true, would be an approximate version of the results of [76, 81, 64] in the hypercube.

Conjecture 2.1.6. *For all $p \in (1/2, 1]$ and $\eta > 0$, a.a.s. $\mathcal{Q}^n_p$ contains $(1/2 - \eta)\delta(\mathcal{Q}^n_p)$ edge-disjoint Hamilton cycles.*

2.1.5 Randomly perturbed graphs

A relatively recent area at the interface of extremal combinatorics and random graph theory is the study of *randomly perturbed graphs*. Generally speaking, the idea is to consider a deterministic dense n-vertex graph H (usually satisfying some minimum degree condition) and a random graph $G_{n,p}$ on the same vertex set as H. The question is whether H is close to satisfying some given property $\mathcal{P}$ in the sense that a.a.s. $H \cup G_{n,p} \in \mathcal{P}$ for some small p. This line of research was sparked off by Bohman, Frieze, and Martin [11], who showed that, if H is an n-vertex graph with $\delta(H) \geq \alpha n$, for any constant $\alpha > 0$, then a.a.s. $H \cup G_{n,p}$ is Hamiltonian for all $p \geq C(\alpha)/n$. Other properties that have been studied in this context are e.g. the existence of powers of Hamilton cycles and general bounded degree spanning graphs [22], F-factors [7] or spanning bounded degree trees [74, 20]. One common phenomenon in this model is that, by considering the union with a dense graph H (i.e. a graph H with linear degrees), the probability threshold of different properties is significantly lower than that in the classical $G_{n,p}$ model. The results for Hamiltonicity [11] were very recently generalised by Hahn-Klimroth, Maesaka, Mogge, Mohr, and Parczyk [53] to allow α to tend to 0 with n (that is, to allow graphs H which are not dense).

We consider randomly perturbed graphs in the setting of subgraphs of the hypercube.

To be precise, we take an arbitrary spanning subgraph H of the hypercube, with linear minimum degree, and a random subgraph $\mathcal{Q}^n_\varepsilon$, and consider $H \cup \mathcal{Q}^n_\varepsilon$. (Note here that $\mathcal{Q}^n_\varepsilon$ is a 'dense' subgraph of $\mathcal{Q}^n$, but for $\varepsilon < 1/2$ it will contain both isolated vertices and vertices of very low degrees.) In this setting, we show the following result.

Theorem 2.1.7. *For all $\varepsilon, \alpha \in (0, 1]$ and $k \in \mathbb{N}$, the following holds. Let H be a spanning subgraph of $\mathcal{Q}^n$ such that $\delta(H) \geq \alpha n$. Then, a.a.s. $H \cup \mathcal{Q}^n_\varepsilon$ contains k edge-disjoint Hamilton cycles.*

We can also allow H to have much smaller degrees, at the cost of requiring a larger probability to find the Hamilton cycles.

Theorem 2.1.8. *For every integer $k \geq 2$, there exists $\varepsilon > 0$ such that a.a.s., for every spanning subgraph H of $\mathcal{Q}^n$ with $\delta(H) \geq k$, the graph $H \cup \mathcal{Q}^n_{1/2-\varepsilon}$ contains a collection of $\lfloor k/2 \rfloor$ Hamilton cycles and $k - 2\lfloor k/2 \rfloor$ perfect matchings, all pairwise edge-disjoint.*

We omit the proof of Theorem 2.1.8 due to word constraints. A full proof can be found in [31]. Note that Theorem 2.1.8 can be viewed as a 'universality' result for H, meaning that it holds for all choices of H simultaneously. It would be interesting to know whether such a result can also be obtained for the lower edge probability assumed in Theorem 3.1.2, i.e., is it the case that, for all $\varepsilon, \alpha \in (0, 1]$, a.a.s. $G \sim \mathcal{Q}^n_\varepsilon$ has the property that, for every spanning $H \subseteq \mathcal{Q}^n$ with $\delta(H) \geq \alpha n$, $G \cup H$ is Hamiltonian?

As we will prove, Theorem 2.1.1 follows straightforwardly from Theorem 2.1.7, and it follows trivially from Theorem 2.1.5. In turn, Theorem 2.1.5 follows from Theorem 2.1.8. On the other hand, Theorems 2.1.2, 2.1.7 and 2.1.8, while being proved with similar ideas, are incomparable.

2.1.6 Percolation on the hypercube

To build Hamilton cycles in random subgraphs of the hypercube, we will consider a random process which can be viewed as a branching process or percolation process on the hypercube.

With high probability, for constant $p > 0$, this process results in a bounded degree tree in $\mathcal{Q}^n_p$ which covers most of the neighbourhood of every vertex in $\mathcal{Q}^n$, and thus spans almost all vertices of $\mathcal{Q}^n$. The version stated below is a special case of Theorem 2.6.1.

Theorem 2.1.9. *For any fixed $\varepsilon, p \in (0,1]$, there exists $D = D(\varepsilon)$ such that a.a.s. $\mathcal{Q}^n_p$ contains a tree T with $\Delta(T) \leq D$ and such that $|V(T) \cap N_{\mathcal{Q}^n}(x)| \geq (1-\varepsilon)n$ for every $x \in V(\mathcal{Q}^n)$.*

Further results concerning the local geometry of the giant component in $\mathcal{Q}^n_p$ for constant $p \in (0, 1/2)$ were proved recently by McDiarmid, Scott, and Withers [83].

The random process we consider in the proof of Theorem 2.1.9 can be viewed as a branching random walk (with a bounded number of branchings at each step). Simpler versions of such processes (with infinite branchings allowed) have been studied by Fill and Pemantle [45] and Kohayakawa, Kreuter, and Osthus [66], and we will base our analysis on these. Motivated by our approach, we raise the following question, which seems interesting in its own right.

Question 2.1.10. *Does a non-returning random walk on $\mathcal{Q}^n$ a.a.s. visit almost all vertices of $\mathcal{Q}^n$?*

More generally, there are many results and applications concerning random walks on the hypercube (but allowing for returns). For example, motivated by a processor allocation problem, Bhatt and Cai [9] studied a walk algorithm to embed large (subdivided) trees into the hypercube. Moreover, the analysis of (branching) random walks is a critical ingredient in the study of percolation thresholds for the existence of a giant component in $\mathcal{Q}^n_p$. These have been investigated e.g. by Bollobás, Kohayakawa, and Łuczak [12] and Borgs, Chayes, Hofstad, Slade, and Spencer [19] and Hofstad and Nachmias [56].

2.2 Outline of the main proofs

2.2.1 Overall outline

We now sketch the key ideas for the proof of Theorem 2.1.7. We will first prove the case $k = 1$, and later use this to deduce the case when $k > 1$. Recall we are given $H \subseteq \mathcal{Q}^n$ with $\delta(H) \geq \alpha n$, and $G \sim \mathcal{Q}^n_\varepsilon$, with $\alpha, \varepsilon \in (0, 1]$. Our aim is to show that a.a.s. $H \cup G$ is Hamiltonian.

Our approach for finding a Hamilton cycle is to first obtain a spanning tree. By passing along all the edges of a spanning tree T (with a vertex ordering prescribed by a depth first search), one can create a closed spanning walk W which visits every edge of T twice. The idea is then to modify such a walk into a Hamilton cycle. (This approach is inspired by the approximation algorithm for the Travelling Salesman Problem which returns a tour of at most twice the optimal length.) More precisely, our approach will be to obtain a near-spanning tree of $\mathcal{Q}^{n-s}$, for some suitable constant s, and to blow up vertices of this tree into s-dimensional cubes. These cubes can then be used to move along the tree without revisiting vertices, which will result in a near-Hamilton cycle $\mathfrak{H}$. All remaining vertices which are not included in $\mathfrak{H}$ will be absorbed into $\mathfrak{H}$ via absorbing structures that we carefully put in place beforehand.

In Sections 2.2.2 to 2.2.4 we outline in more detail how we find a long cycle in G (Theorem 2.1.2). Note that in Theorem 2.1.2 we have $G \sim \mathcal{Q}^n_\varepsilon$, so a.a.s. G will have isolated vertices which prevent any Hamilton cycle occurring as a subgraph. In Section 2.2.5 we outline how we build on this approach to obtain the case $k = 1$ of Theorem 2.1.7. In Section 2.2.6 we sketch how we obtain Theorem 2.1.5.

2.2.2 Building block I: trees via branching processes.

We view each vertex in $\mathcal{Q}^n$ as an n-dimensional 01-coordinate vector. By fixing the first s coordinates, we fix one of 2^s *layers* $L_1, \dots, L_{2^s}$ of the hypercube, where $s \in \mathbb{N}$ will be constant. Thus, $L \cong \mathcal{Q}^{n-s}$ for each layer L. By considering a Hamilton cycle in $\mathcal{Q}^s$, we may assume that consecutive layers differ only by a single coordinate on the unique elements of $\mathcal{Q}^s$ which define them. Let $G \sim \mathcal{Q}^n_\varepsilon$. For each layer L, we let $L(G) \coloneqq G[V(L)]$ and define the *intersection graph* $I(G) \coloneqq \bigcap_{i=1}^{2^s} L_i(G)$. Hence, $I(G) \sim \mathcal{Q}^{n-s}_{\varepsilon^{2^s}}$. We view $I(G)$ as a subgraph of $\mathcal{Q}^{n-s}$. We first show that $I(G)$ contains a near-spanning tree T (Theorem 2.6.1). Thus, a copy of T is present in each of $L_1(G), \dots, L_{2^s}(G)$ simultaneously.

Since the walk W mentioned in Section 2.2.1 passes through each vertex x of T a total of $d_T(x)$ times, it will be important later for T to have bounded degree. In order to guarantee this, we run bounded degree branching processes (see Definition 2.6.3) from several far apart 'corners' of the hypercube. Roughly speaking, T will be formed by taking a union of these processes and removing cycles. Crucially, the model we introduce for these processes has a joint distribution with $\mathcal{Q}^{n-s}_{\varepsilon^{2^s}}$, so that T will in fact appear as a subgraph of $I(G)$. In applying Theorem 2.6.1, we obtain a bounded degree tree $T \subseteq I(G)$ which contains almost all of the neighbours of every vertex of $I(G)$. We also obtain a 'small' *reservoir* set $R \subseteq V(I(G))$, which T avoids and which will play a key role later in the absorption of vertices which do not belong to our initial long cycle. At this point, both T and R are now present in every layer of the hypercube simultaneously.

2.2.3 Building block II: cube tilings via the nibble.

Let $\ell < s$ and $0 < \delta \ll 1$ be fixed. In order to gain more local flexibility when traversing the near-spanning tree T, we augment T by locally adding a near-spanning ℓ-cube factor of $I(G)$. One can use classical results on matchings in almost regular uniform hypergraphs of small codegree to show that $I(G)$ contains such a collection of $\mathcal{Q}^\ell$ spanning almost all vertices of $I(G)$. However, we require the following stronger properties, namely that there exists a collection $\mathcal{C}$ of vertex disjoint copies of $\mathcal{Q}^\ell$ in $I(G)$ so that, for each $x \in V(I(G))$,

(i) $\mathcal{C}$ covers almost all vertices in $N_{\mathcal{Q}^n}(x)$;

(ii) the directions spanned by the cubes intersecting $N_{\mathcal{Q}^n}(x)$ do not correlate too strongly with any given set of directions.

The precise statement is given in Theorem 2.5.7. Neither (i) nor (ii) follow from existing results on hypergraph matchings and the proofs strongly rely on geometric properties intrinsic to the hypercube. We will omit the proof of Theorem 2.5.7 due to word constraints but describe the main steps of this proof below.

To prove Theorem 2.5.7, we build on the so-called Rödl nibble. More precisely, we consider the hypergraph $\mathcal{H}$, with $V(\mathcal{H}) = V(\mathcal{Q}^{n-s})$, where the edge set is given by the copies of $\mathcal{Q}^\ell$ in $I(G)$. We run a random iterative process where at each stage we add a 'small' number of edges from $\mathcal{H}$ to $\mathcal{C}$, before removing all those remaining edges of $\mathcal{H}$ which 'clash' with our selection. A careful analysis and an application of the Lovász local lemma yield the existence of an instance of this process which terminates in the near-spanning ℓ-cube factor with the properties required for Theorem 2.5.7.

2.2.4 Constructing a long cycle.

Roughly speaking, we will use T as a backbone to provide 'global' connectivity, and will use the near-spanning ℓ-cube factor $\mathcal{C}$ and the layer structure to gain high 'local' connectivity and flexibility. Let $T \cup \bigcup_{C \in \mathcal{C}} C =: \Gamma' \subseteq I(G)$ and let $\Gamma \subseteq \Gamma'$ be formed by removing all leaves and isolated cubes in Γ'. It follows by our tree and nibble results that almost all vertices of $I(G)$ are contained in Γ. Note that, for each $v \in V(\mathcal{Q}^{n-s}) = V(I(G))$, there is a unique vertex in each of the 2^s layers which corresponds to v. We refer to these 2^s vertices as *clones* of v and to the collection of these 2^s clones as a *vertex molecule*. Similarly, each ℓ-cube $C \in \mathcal{C}$

contained in Γ gives rise to a *cube molecule*. We construct a cycle in G which covers all of the cube molecules (and, therefore, almost all vertices in $\mathcal{Q}^n$).

Let Γ^* be the graph obtained from Γ by contracting each ℓ-cube $C \subseteq (\bigcup_{C \in \mathcal{C}} C) \cap \Gamma$ into a single vertex. We refer to such vertices in Γ^* as *atomic vertices*, and to all other vertices as *inner tree vertices*. We run a depth-first search on Γ^* to give an order to the vertices. Next, we construct a *skeleton* which will be the backbone for our long cycle. The skeleton is an ordered sequence of vertices in $\mathcal{Q}^n$ which contains the vertices via which our cycle will enter and exit each molecule. That is, given an *exit vertex* v for some molecule in the skeleton, the vertex u which succeeds v in the skeleton will be an *entry vertex* for another molecule, and such that $uv \in E(G)$. Here, a vertex in the skeleton belonging to an inner tree vertex molecule is referred to as both an entry and exit vertex. (Actually, we will first construct an 'external skeleton', which encodes this information. The skeleton then also prescribes some edges within molecules which go between different layers.) We use the ordering of the vertices of Γ^* to construct the skeleton in a recursive way starting from the lowest ordered vertex. It is crucial that our tree T has bounded degree (much smaller than 2^s), so that no molecule is overused in the skeleton.

Once the skeleton is constructed, we apply our 'connecting lemmas' (Lemmas 2.7.8 and 2.7.9). These connecting lemmas, applied to a cube molecule with a bounded number of pairs of entry and exit vertices as input (given by the skeleton), provide us with a sequence of vertex-disjoint paths which cover this molecule, where each path has start and end vertices consisting of an input pair. The union of all of these paths combined with all edges in G between the successive exit and entry vertices of the skeleton will then form a cycle $\mathfrak{H} \subseteq G$ which covers all vertices lying in the cube molecules (thus proving Theorem 2.1.2).

2.2.5 Constructing a Hamilton cycle.

In order to construct a Hamilton cycle in $H \cup G$, we will absorb the vertices of $V(\mathcal{Q}^n) \setminus V(\mathfrak{H})$ into $\mathfrak{H}$. We achieve this via absorbing structures that we identify for each vertex (see Definition 2.7.2). To construct these absorbing structures, we will need to use some edges of H. Roughly speaking, to each vertex v we associate a left ℓ-cube $C_v^l \subseteq \mathcal{Q}^n$ and a right ℓ-cube $C_v^r \subseteq \mathcal{Q}^n$, where C_v^l, C_v^r are both clones of some ℓ-cubes $C^l, C^r \in \mathcal{C}$ contained in Γ. We choose these cubes so that v will have a neighbour $u \in V(C_v^l)$ and a neighbour $u' \in V(C_v^r)$, to which we refer as *tips* of the absorbing structure. Furthermore, u will have a neighbour $w \in V(C_v^r)$, which is also a neighbour of u'. Our near-Hamilton cycle $\mathfrak{H}$ will satisfy the following properties:

(a) $\mathfrak{H}$ covers all vertices in $C_v^l \cup C_v^r$ except for u, and

(b) $wu' \in E(\mathfrak{H})$.

These additional properties will be guaranteed by our connecting lemmas discussed in Section 2.2.4. We can then alter $\mathfrak{H}$ to include the segment $wuvu'$ instead of the edge wu', thus absorbing the vertices u and v into $\mathfrak{H}$. The following types of vertices will require absorption.

(i) Every vertex that is not covered by a clone of either some inner tree vertex or of some cube $C \in \mathcal{C}$ which is contained in Γ.

(ii) The cycle $\mathfrak{H}$ does not cover all the clones of inner tree vertices and, thus, the uncovered vertices of this type will also have to be absorbed.

However, we will not know precisely which of the vertices described in (i) and (ii) will be covered by $\mathfrak{H}$ and which of these vertices will need to be absorbed until after we have constructed the (external) skeleton. Moreover, many potential absorbing structures are later ruled out as candidates (for example, if they themselves contain vertices that will need to be absorbed). Therefore, it is important that we identify a 'robust' collection of many potential absorbing structures for every vertex in $\mathcal{Q}^n$ at a preliminary stage of the proof. The precise absorbing structure eventually assigned to each vertex will be chosen via an application of our rainbow matching lemma (Lemma 2.5.4) at a late stage in the proof.

We will now highlight the purpose of the reservoir R. Suppose $v \in V(\mathcal{Q}^n)$ is a vertex which needs to be absorbed via an absorbing structure with left ℓ-cube C_v^l and left tip $u \in V(C_v^l)$. Recall that both u and C_v^l are clones of some $u^* \in V(\Gamma)$ and $C^l \in \mathcal{C}$, where $u^* \in V(C^l)$. If u^* has a neighbour w^* in $T - V(C^l)$, then it is possible that the skeleton will assign an edge from u to w for the cycle $\mathfrak{H}$ (where w is the clone of w^* in the same layer as u). Given that u is now incident to a vertex outside of C_v^l, we can no longer use the absorbing structure with u as a (left) tip (otherwise, we might disconnect T). To avoid this problem, we show that most vertices have many potential absorbing structures whose tips lie in the reservoir R (which T avoids). Here we make use of vertex degrees of H. A small number of *scant vertices* will not have high enough degree into R. For these vertices we fix an absorbing structure whose tips do not lie in R, and then alter T slightly so that these tips are deleted from T and reassigned to R. The fact that scant vertices are few and well spread out from each other will be crucial in being able to achieve this (see Lemma 2.6.20).

Let us now discuss two problems arising in the construction of the skeleton. Firstly, let $\mathcal{M}_C \subseteq \mathcal{Q}^n$ with $C \in \mathcal{C}$ be a cube molecule which is to be covered by $\mathfrak{H}$. Furthermore, suppose one of the clones C_v^l of C belongs to an absorbing structure for some vertex v. Let u be the tip of C_v^l and suppose that u has even parity. We would like to apply the connecting lemmas to cover $\mathcal{M}_C - \{u\}$ by paths which avoid u. But this would now involve covering one fewer vertex of even parity than of odd parity. This, in turn, has the effect of making the construction of the skeleton considerably more complicated (this construction is simplest when successive entry and exit vertices have opposite parities). To avoid this, we assign absorbing structures in pairs, so that, for each $C \in \mathcal{C}$, either two or no clones of C will be used in absorbing structures. In the case where two clones are used, we enforce that the tips of these clones will have opposite parities, and therefore each molecule $\mathcal{M}_C$ will have the same number of even and odd parity vertices to be covered by $\mathfrak{H}$. We use our robust matching lemma (see Lemma 2.5.2) to pair up the clones of absorbing structures in this way. To connect up different layers of a cube molecule, we will of course need to have suitable edges between these. Molecules which do not satisfy this requirement are called 'bondless' and are removed from Γ before the absorption process (so that their vertices are absorbed).

Secondly, another issue related to vertex parities arises from inner tree vertex molecules. Depending on the degree of an inner tree vertex $v \in V(T)$, the skeleton could contain an odd number of vertices from the molecule $\mathcal{M}_v$ consisting of all clones of v. All vertices in $\mathcal{M}_v$ outside the skeleton will need to be absorbed. But since the number of these vertices is odd, it would be impossible to pair up (in the way described above) the absorbing structures assigned to these vertices. To fix this issue, we effectively impose that $\mathfrak{H}$ will 'go around T twice'. That is, the skeleton will trace through every molecule beginning and finishing at the lowest ordered vertex in Γ^*. It will then retrace its steps through these molecules in an almost identical way, effectively doubling the size of the skeleton. This ensures that the skeleton contains an even number of vertices from each molecule, half of them of each parity.

Finally, once we have obtained an appropriate skeleton, we can construct a long cycle $\mathfrak{H}$ as described in Section 2.2.4. For every vertex in $\mathcal{Q}^n$ which is not covered by $\mathfrak{H}$ we have put in place an absorbing structure, which is covered by $\mathfrak{H}$ as described in (a) and (b). Thus, as discussed before, we can now use these structures to absorb all remaining vertices into $\mathfrak{H}$ to obtain a Hamilton cycle $\mathfrak{H}' \subseteq H \cup G$, thus proving the case $k = 1$ of Theorem 2.1.7.

2.2.6 Hitting time for the appearance of a Hamilton cycle.

As mentioned in Section 2.1.4, we will omit the proof of Theorem 2.1.5 due to word constraints. We offer the following insight into the proof by highlighting the key steps where it builds on the proof of Theorem 2.7.1.

In order to prove Theorem 2.1.5, we consider $G \sim \mathcal{Q}^n_{1/2-\varepsilon}$. We show that a.a.s., for any graph H with $\delta(H) \geq 2$, the graph $G \cup H$ is Hamiltonian. The main additional difficulty faced here is that $G \cup H$ may contain vertices having degree as low as 2. For the set $\mathcal{U}$ of these vertices we cannot hope to use the previous absorption strategy: the neighbours of $v \in \mathcal{U}$ may not lie in cubes from $\mathcal{C}$. (In fact, v may not even have a neighbour within its own layer in $G \cup H$.) To handle such small degree vertices, we first prove that they will be few and well spread out. We define three types of new 'special absorbing structures'. The type of the special absorbing structure $SA(v)$ for v will depend on whether the neighbours a, b of v in H lie in the same layer as v. In each case, $SA(v)$ will consist of a short path P_1 containing the edges av and bv, and several other short paths designed to 'balance out' P_1 in a suitable way. These paths will be incorporated into the long cycle $\mathfrak{H}$ described in Section 2.2.4. In particular, this allows us to 'absorb' the vertices of $\mathcal{U}$ into $\mathfrak{H}$. To incorporate the paths P_i forming $SA(v)$, we will proceed as follows.

Firstly, we make use of the fact that Theorem 2.6.1 allows us to choose our near-spanning tree T in such a way that it avoids a small ball around each $v \in \mathcal{U}$. Thus, (all clones of) T will avoid $SA(v)$, which has the advantage there will be no interference between T and the special absorbing structures. To link up each $SA(v)$ with the long cycle $\mathfrak{H}$, for each endpoint w of a path in $SA(v)$, we will choose an ℓ-cube in $I(G)$ which suitably intersects T and which contains w (or more precisely, the vertex in $I(G)$ corresponding to w). Altogether, these ℓ-cubes allow us to find paths between $SA(v)$ and vertices of $\mathfrak{H}$ which are clones of vertices in T. The remaining vertices in molecules consisting of clones of these ℓ-cubes will be covered in a similar way as in Section 2.2.4. All vertices in these balls around $\mathcal{U}$ which are not part of the special absorbing structures will be absorbed into $\mathfrak{H}$ via the same absorbing structures used in the proof of Theorem 2.1.7 to once again obtain a Hamilton cycle $\mathfrak{H}'$.

2.2.7 Edge-disjoint Hamilton cycles.

The results on k edge-disjoint Hamilton cycles can be deduced from suitable versions of the case $k = 1$. Those versions are carefully formulated to allow us to repeatedly remove a Hamilton cycle from the original graph. We deduce Theorem 2.1.1 from Theorem 2.7.1 in Section 2.7.5.

2.3 Notation

For $n \in \mathbb{Z}$, we denote $[n] := \{k \in \mathbb{Z} : 1 \leq k \leq n\}$ and $[n]_0 := \{k \in \mathbb{Z} : 0 \leq k \leq n\}$. Whenever we write a hierarchy of parameters, these are chosen from right to left. That is, whenever we claim that a result holds for $0 < a \ll b \leq 1$, we mean that there exists a non-decreasing function $f\colon [0,1) \to [0,1)$ such that the result holds for all $a > 0$ and all $b \leq 1$ with $a \leq f(b)$. We will not compute these functions explicitly. Hierarchies with more constants are defined in a similar way.

A *hypergraph* H is an ordered pair $H = (V(H), E(H))$ where $V(H)$ is called the vertex set and $E(H) \subseteq 2^{V(H)}$, the edge set, is a set of subsets of $V(H)$. If $E(H)$ is a multiset, we refer to H as a *multihypergraph*. We say that a (multi)hypergraph H is r-*uniform* if for every $e \in E(H)$ we have $|e| = r$. In particular, 2-uniform hypergraphs are simply called *graphs*. Given any set of vertices $V' \subseteq V(H)$, we denote the subhypergraph of H *induced* by V' as $H[V'] := (V', E')$, where $E' := \{e \in E(H) : e \subseteq V'\}$. We write $H - V' := H[V \setminus V']$. Given any set $\hat{E} \subseteq E(H)$, we will sometimes write $V(\hat{E}) := \{v \in V :$ there exists $e \in \hat{E}$ such that $v \in e\}$.

Given any (multi)hypergraph H and any vertex $v \in V(H)$, let $E(H, v) := \{e \in E(H) : v \in e\}$. We define the *neighbourhood* of v as $N_H(v) := \bigcup_{e \in E(H,v)} e \setminus \{v\}$, and we define the *degree* of v by $d_H(v) := |E(H, v)|$. We denote the minimum and maximum degrees of (the vertices in) H by $\delta(H)$ and $\Delta(H)$, respectively. Given any pair of vertices $u, v \in V(H)$, we define $E(H, u, v) := \{e \in E(H) : \{u, v\} \subseteq e\}$. The *codegree* of u and v

in H is given by $d_H(u,v) \coloneqq |E(H,u,v)|$. Given any set of vertices $W \subseteq V(H)$, we define $N_H(W) \coloneqq \bigcup_{w\in W} N_H(w)$. We denote $E(H,v,W) \coloneqq \{e \in E(H) : v \in e, e \setminus \{v\} \subseteq W\}$, $N_H(v,W) \coloneqq \bigcup_{e\in E(H,v,W)} e \setminus \{v\}$ and $d_H(v,W) \coloneqq |E(H,v,W)|$; we refer to the latter two as the neighbourhood and degree of v into W, respectively. Given $A, B \subseteq V(H)$ we denote $E_H(A,B) \coloneqq \{e \in E(H) : e \subseteq A \cup B, e \cap A \neq \varnothing, e \cap B \neq \varnothing\}$ and $e_H(A,B) \coloneqq |E_H(A,B)|$. Whenever $A = \{v\}$ is a singleton, we abuse notation and write $E_H(v,B)$ and $e_H(v,B)$. Thus, $e_H(v,B)$ and $d_H(v,B)$ may be used interchangeably.

Given any graph G and two vertices $u, v \in V(G)$, the *distance* $\mathrm{dist}_G(u,v)$ between u and v in G is defined as the length of the shortest path connecting u and v (and it is said to be infinite if there is no such path). Similarly, given any sets $A, B \subseteq V(G)$, the *distance* between A and B is given by $\mathrm{dist}_G(A,B) \coloneqq \min_{u\in A, v\in B} \mathrm{dist}_G(u,v)$. For any $r \in \mathbb{N}$, we denote $B^r_G(u) \coloneqq \{v \in V(G) : \mathrm{dist}_G(u,v) \leq r\}$ and $B^r_G(A) \coloneqq \{v \in V(G) : \mathrm{dist}_G(A,v) \leq r\}$; we refer to these sets as the *balls* of radius r around u and A, respectively.

A *directed graph* (or *digraph*) is a pair $D = (V(D), E(D))$, where $E(D)$ is a set of ordered pairs of elements of $V(D)$. If no pair of the form (v,v) with $v \in V(D)$ belongs to $E(D)$, we say that D is *loopless*. Given any $v \in V(D)$, we define its *inneighbourhood* as $N^-_D(v) \coloneqq \{u \in V(D) : (u,v) \in E(D)\}$, and its *outneighbourhood* as $N^+_D(v) \coloneqq \{u \in V(D) : (v,u) \in E(D)\}$. The *indegree* and *outdegree* of v are defined as $d^-_D(x) \coloneqq |N^-_D(x)|$ and $d^+_D(x) \coloneqq |N^+_D(x)|$, respectively. The minimum in- and outdegrees of (the vertices in) D are denoted by $\delta^-(D)$ and $\delta^+(D)$, respectively.

Given any multihypergraph or directed graph (V,E), a set $M \subseteq E$ is called a *matching* if its elements are pairwise disjoint. If the edges of M cover all of V, then it is said to be a *perfect matching*. Given an edge-colouring c of H, we say that a matching of H is *rainbow* if each of its edges has a different colour in c.

We often refer to the n-dimensional hypercube $\mathcal{Q}^n$ as an *n-cube* (the n is dropped whenever clear from the context). Given two vertices $v_1, v_2 \in V(\mathcal{Q}^n) = \{0,1\}^n$, we write $\mathrm{dist}(v_1,v_2)$ for the Hamming distance between v_1 and v_2. Thus, $\{v_1,v_2\} \in E(\mathcal{Q}^n)$ if and only if $\mathrm{dist}(v_1, v_2) = 1$. Whenever the dimension n is clear from the context, we will use $\mathbf{0}$ to denote the vertex $\{0\}^n$. Given any $v \in \{0,1\}^n$, we will say that its *parity* is *even* if $\mathrm{dist}(v, \mathbf{0}) \equiv 0 \pmod 2$, and we will say that it is *odd* otherwise. This gives a natural partition of $V(\mathcal{Q}^n)$ into the sets of vertices with even and odd parities. Given any two vertices $v_1, v_2 \in \{0,1\}^n$, we will write $v_1 =_{\mathrm{p}} v_2$ if they have the same parity, and $v_1 \neq_{\mathrm{p}} v_2$ otherwise.

We will often consider the natural embedding of $V(\mathcal{Q}^n)$ into $\mathbb{F}_2^n$, which will allow us to use operations on the vertex set: whenever we write $v + u$, for some $u, v \in \{0,1\}^n$, we refer to their sum in $\mathbb{F}_2^n$. Given a vertex $v \in \{0,1\}^n$ and an edge $e = \{x, y\} \in E(\mathcal{Q}^n)$, we define $v + e$ to be the edge with endvertices $v + x$ and $v + y$. Given any two sets $A, B \subseteq \{0,1\}^n$, we will use the sumset notation $A + B \coloneqq \{a + b : a \in A, b \in B\}$, and we will abbreviate the k-fold sumset $A + \ldots + A$ by kA. Similarly, given any sets $A \subseteq \{0,1\}^n$ and $E \subseteq E(\mathcal{Q}^n)$, we write $A + E \coloneqq \{a + e : a \in A, e \in E\}$. Given a graph $G \subseteq \mathcal{Q}^n$ and a set of vertices $A \subseteq \{0,1\}^n$, $A + G$ will denote the graph with vertex set $A + V(G)$ and edge set $A + E(G)$. Note that this should never be confused with the notation $G - A$, which will be used exclusively to consider induced subgraphs of G. We will call the unitary vectors in $\mathbb{F}_2^n$ the *directions* of the hypercube. The set of directions will be denoted by $\mathcal{D}(\mathcal{Q}^n)$. Thus, $\mathcal{D}(\mathcal{Q}^n) = \{\hat{e} \in \{0,1\}^n : \mathrm{dist}(\hat{e}, \mathbf{0}) = 1\}$. Note that two vertices $v_1, v_2 \in \{0,1\}^n$ are adjacent in $\mathcal{Q}^n$ if and only if there exists $\hat{e} \in \mathcal{D}(\mathcal{Q}^n)$ such that $v_1 = v_2 + \hat{e}$. Given any vertex $v \in \{0,1\}^n$ and any set $\mathcal{D} \subseteq \mathcal{D}(\mathcal{Q}^n)$, we will denote by $\mathcal{Q}^n(v, \mathcal{D}) \coloneqq \mathcal{Q}^n[v + n(\mathcal{D} \cup \{\mathbf{0}\})]$ the subcube of $\mathcal{Q}^n$ which contains v and all vertices in $\{0,1\}^n$ which can be reached from v by only adding directions in $\mathcal{D}$. Given any subcube $Q \subseteq \mathcal{Q}^n$, we will write $\mathcal{D}(Q)$ to denote the subset of $\mathcal{D}(\mathcal{Q}^n)$ such that, for any $v \in V(Q)$, we have $Q = \mathcal{Q}^n(v, \mathcal{D}(Q))$. Given any direction $\hat{e} \in \mathcal{D}(Q)$, we will sometimes informally say that Q *uses* $\hat{e}$. Given two vertices $v_1, v_2 \in \{0,1\}^n$, their *differing directions* are all directions in $\mathcal{D}(v_1, v_2) \coloneqq \{\hat{e} \in \mathcal{D}(\mathcal{Q}^n) : \mathrm{dist}(v_1 + \hat{e}, v_2) < \mathrm{dist}(v_1, v_2)\}$. Observe that, if $\mathrm{dist}(v_1, v_2) = d$, then $|\mathcal{D}(v_1, v_2)| = d$ and $\mathcal{Q}^n(v_1, \mathcal{D}(v_1, v_2))$ is the smallest subcube of $\mathcal{Q}^n$ which contains both v_1 and v_2.

When considering random experiments for a sequence of graphs $(G_n)_{n \in \mathbb{N}}$ with $|V(G_n)|$

tending to infinity with n, we say that an event $\mathcal{E}$ holds *asymptotically almost surely* (*a.a.s.*) for G_n if $\mathbb{P}[\mathcal{E}] = 1 - o(1)$. When considering asymptotic statements, we will ignore rounding whenever this does not affect the argument.

2.4 Probabilistic tools

Here we list some probabilistic tools that we will use throughout the paper. The following can be proved easily with the Cauchy-Schwarz inequality.

Proposition 2.4.1. *Given a non-negative random variable X with finite support, we have that*

$$\mathbb{P}[X = 0] \leq 1 - \frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2]}.$$

Throughout the paper, we will be interested in proving concentration results for different random variables. We will often need the following Chernoff bound (see e.g. [58, Corollary 2.3]).

Lemma 2.4.2. *Let X be the sum of n mutually independent Bernoulli random variables and let $\mu := \mathbb{E}[X]$. Then, for all $0 < \delta < 1$ we have that $\mathbb{P}[X \geq (1+\delta)\mu] \leq e^{-\delta^2\mu/3}$ and $\mathbb{P}[X \leq (1-\delta)\mu] \leq e^{-\delta^2\mu/2}$. In particular, $\mathbb{P}[|X - \mu| \geq \delta\mu] \leq 2e^{-\delta^2\mu/3}$.*

Similar bounds hold for hypergeometric distributions (see e.g. [58, Theorem 2.10]). For $m, n, N \in \mathbb{N}$ with $m, n < N$, a random variable X is said to follow the hypergeometric distribution with parameters N, n and m if it can be defined as $X := |S \cap [m]|$, where S is a uniformly chosen random subset of $[N]$ of size n.

Lemma 2.4.3. *Suppose Y has a hypergeometric distribution with parameters N, n and m. Then, $\mathbb{P}[|Y - \mathbb{E}[Y]| \geq t] \leq 2e^{-t^2/(3n)}$.*

The following bound will also be used repeatedly (see e.g. [5, Theorem A.1.12]).

Lemma 2.4.4. *Let X be the sum of n mutually independent Bernoulli random variables. Let $\mu := \mathbb{E}[X]$, and let $\beta > 1$. Then, $\mathbb{P}[X \geq \beta\mu] \leq (e/\beta)^{\beta\mu}$. In particular, we have $\mathbb{P}[X \geq 7\mu] \leq e^{-\mu}$.*

Finally, the Lovász local lemma will come in useful. Let $\mathfrak{E} := \{\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_m\}$ be a collection of events. A *dependency graph* for $\mathfrak{E}$ is a graph H on vertex set $[m]$ such that, for all $i \in [m]$, $\mathcal{E}_i$ is mutually independent of $\{\mathcal{E}_j : j \neq i, j \notin N_H(i)\}$, that is, if $\mathbb{P}[\mathcal{E}_i] = \mathbb{P}[\mathcal{E}_i \mid \bigwedge_{j\in J} \mathcal{E}_j]$ for all $J \subseteq [m] \setminus (N_H(i) \cup \{i\})$. We will use the following version of the local lemma (it follows e.g. from [5, Lemma 5.1.1]).

Lemma 2.4.5 (Lovász local lemma)**.** *Let $\mathfrak{E} := \{\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_m\}$ be a collection of events and let H be a dependency graph for $\mathfrak{E}$. Suppose that $\Delta(H) \leq d$ and $\mathbb{P}[\mathcal{E}_i] \leq p$ for all $i \in [m]$. If $ep(d+1) \leq 1$, then*

$$\mathbb{P}\left[\bigwedge_{i=1}^{m} \overline{\mathcal{E}_i}\right] \geq (1-ep)^m.$$

2.5 Auxiliary results

2.5.1 Results about matchings

We will need three auxiliary results to help us find suitable absorbing cube pairs for different vertices. We will need to preserve the alternating parities of vertices that are absorbed by each molecule. The first lemma (Lemma 2.5.2) presented in this section, will help us to show that all vertices can be paired up in such a way that these parities can be preserved. The second lemma (Lemma 2.5.3) will be used to show that, for each such pair of vertices, there are many possible pairs of absorption cubes. Finally, the third lemma (Lemma 2.5.4) will allow us to assign one of those pairs of absorption cubes to each pair of vertices we need to absorb in such a way that these cube pairs are pairwise vertex disjoint.

To prove Lemma 2.5.2, as well as Lemma 2.6.16 and Theorem 2.6.19, the following consequence of Hall's theorem will be useful.

Lemma 2.5.1. *Let G be a bipartite graph with vertex partition $A \dot{\cup} B$. Assume that there is some integer $\ell \geq 0$ such that, for all $S \subseteq A$, we have $|N(S)| \geq |S| - \ell$. Then, G contains a matching which covers all but at most ℓ vertices in A.*

Given any graph G and a bipartition $(\mathfrak{A}, \mathfrak{B})$ of $V(G)$, we say that $(\mathfrak{A}, \mathfrak{B})$ is an *r-balanced bipartition* if $||\mathfrak{A}| - |\mathfrak{B}|| \leq r$. Let G be a graph on n vertices, and let $r, d \in \mathbb{N}$ with $r \leq d$. We say that G is *d-robust-parity-matchable* with respect to an r-balanced bipartition $(\mathfrak{A}, \mathfrak{B})$ if, for every $S \subseteq V(G)$ such that $|S| \leq d$ and $|\mathfrak{A} \setminus S| = |\mathfrak{B} \setminus S|$, the graph $G - S$ contains a perfect matching M with the property that every edge $e \in M$ has one endpoint in $\mathfrak{A} \setminus S$ and one endpoint in $\mathfrak{B} \setminus S$.

Given two disjoint sets of vertices A and B, the binomial random bipartite graph $G(A, B, p)$ is obtained by adding each possible edge with one endpoint in A and the other in B with probability p independently of every other edge. Given any two bipartite graphs on the same vertex set, $G_1 = (A, B, E_1)$ and $G_2 = (A, B, E_2)$, and any $\alpha \in \mathbb{R}$, we define $\Gamma^{\alpha}_{G_1,G_2}(A)$ as the graph with vertex set A where any two vertices $x, y \in A$ are joined by an edge whenever $|N_{G_1}(x) \cap N_{G_2}(y)| \geq \alpha|B|$ or $|N_{G_1}(y) \cap N_{G_2}(x)| \geq \alpha|B|$.

Lemma 2.5.2. *Let $d, k, r \in \mathbb{N}$ and $\alpha, \varepsilon, \beta > 0$ be such that $r \leq d$, $1/k \ll 1/d, \varepsilon, \alpha$ and $\beta \ll \varepsilon, \alpha$. Then, any bipartite graph $G = G(A, B, E)$ with $|B| = n \geq |A| \geq k$ such that $d_G(x) \geq \alpha n$ for every $x \in A$ satisfies the following with probability at least $1 - 2^{-10n}$: for any r-balanced bipartition of A into $(\mathfrak{A}, \mathfrak{B})$, the graph $\Gamma^{\beta}_{G,G(A,B,\varepsilon)}(A)$ is d-robust-parity-matchable with respect to $(\mathfrak{A}, \mathfrak{B})$.*

Proof. Let $\Gamma \coloneqq \Gamma^{\beta}_{G,G(A,B,\varepsilon)}(A)$. Let Γ' be the auxiliary digraph with vertex set A where, for any pair of vertices $x, y \in A$, there is a directed edge from x to y if $|N_G(x) \cap N_{G(A,B,\varepsilon)}(y)| \geq \beta n$. Observe that the graph obtained from Γ' by ignoring the directions of its edges and identifying the possible multiple edges is exactly Γ, which means that $\delta(\Gamma) \geq \delta^+(\Gamma')$.

Given any two vertices $x, y \in A$, by Lemma 2.4.2 we have that

$$\mathbb{P}[(x, y) \notin E(\Gamma')] = \mathbb{P}[|N_G(x) \cap N_{G(A,B,\varepsilon)}(y)| < \beta n] \leq e^{-\varepsilon\alpha n/3}.$$

Furthermore, for a fixed $x \in A$, observe that the events that $(x, y) \notin E(\Gamma')$, for all $y \in A \setminus \{x\}$, are mutually independent. Therefore, $d^+_{\Gamma'}(x)$ is a sum of independent Bernoulli random variables. Let $m := |A|$. If $d^+_{\Gamma'}(x) < 4m/5$, that means that there is a set of $m/5$ vertices $Y \subseteq A \setminus \{x\}$ such that $(x, y) \notin E(\Gamma')$ for all $y \in Y$. We then conclude that

$$\mathbb{P}[d^+_{\Gamma'}(x) < 4m/5] \leq \sum_{Y \in \binom{A \setminus \{x\}}{m/5}} \mathbb{P}[(x, y) \notin E(\Gamma') \text{ for all } y \in Y] \leq \binom{m}{m/5} e^{-\varepsilon\alpha nm/15} \leq 2^{-20n}.$$

By a union bound over the choice of x, we conclude that

$$\mathbb{P}[\delta(\Gamma) < 4m/5] \leq \mathbb{P}[\delta^+(\Gamma') < 4m/5] \leq m2^{-20n} \leq 2^{-10n}.$$

Now, condition on the event that the previous holds. Fix any r-balanced bipartition $(\mathfrak{A}, \mathfrak{B})$ of A and let $\Gamma_{(\mathfrak{A},\mathfrak{B})}$ be the bipartite subgraph of Γ induced by this bipartition. Fix any set $S \subseteq A$ with $|S| \leq d$ and $|\mathfrak{A} \setminus S| = |\mathfrak{B} \setminus S|$. We have that $\delta(\Gamma_{(\mathfrak{A},\mathfrak{B})} - S) \geq 4m/5 - m/2 - d - r > m/4$. Therefore, by Lemma 2.5.1, $\Gamma_{(\mathfrak{A},\mathfrak{B})} - S$ contains a perfect matching. □

The second lemma will be stated in terms of directed graphs.

Lemma 2.5.3. *Let $c, C > 0$ and let $\alpha \in (0, 1/(1 + c/C))$. Let D be a loopless n-vertex digraph such that*

(i) *for every $A \subseteq V(D)$ with $|A| \geq \alpha n$ we have $\sum_{v \in A} d^-(v) \geq c\alpha n$, and*

(ii) *for every $B \subseteq V(D)$ with $|B| \leq c\alpha n/C$ we have $\sum_{v \in B} d^+(v) \leq c\alpha n$.*

Then, D contains a matching M with $|M| > c\alpha n/(2C)$.

Proof. Assume for a contradiction that the largest matching M in D has size $|M| \leq c\alpha n/(2C)$. Since $\alpha < 1/(1 + c/C)$, there exists a set $A \subseteq V(D) \setminus V(M)$ with $|A| \geq \alpha n$, and thus, by (i), $\sum_{v \in A} d^-(v) \geq c\alpha n$. Since M is the largest matching, all edges that enter A must come from vertices of M (otherwise, we could add one such edge to M, finding a larger matching). However, by (ii), the number of edges going out of $V(M)$ is less than $c\alpha n$, a contradiction. □

For convenience, we state the third lemma in terms of rainbow matchings in hypergraphs.

Lemma 2.5.4. *Let $n, r \in \mathbb{N}$ and let $\mathcal{H}$ be an n-edge-coloured r-uniform multihypergraph. Then, for any $m \geq 10$, the following holds. Suppose $\mathcal{H}$ satisfies the following two properties:*

(i) *For every $i \in [n]$, there are at least m edges of colour i.*

(ii) $\Delta(\mathcal{H}) \leq m/(6r)$.

Then, there exists a rainbow matching of size n.

Proof. The idea is to pick a random edge from each colour class and prove that with non-zero probability this results in a rainbow matching. First, for each $i \in [n]$, let M_i be a set of m edges of colour i. We choose an edge from each M_i uniformly at random, independently of the other choices. For any $i, j \in [n]$ with $i \neq j$ and for any two edges $e \in M_i$ and $e' \in M_j$ for which $e \cap e' \neq \varnothing$, we denote by $A_{e,e'}$ the event that both e and e' are picked. We observe that

$$\mathbb{P}[A_{e,e'}] = \left(\frac{1}{m}\right)^2.$$

Moreover, note that every event $A_{e,e'}$ is independent of all other events $A_{f,f'}$ but at most $2m \cdot r \cdot \Delta(\mathcal{H}) \leq m^2/3$. Indeed, this holds because $A_{e,e'}$ can only depend on those events which involve at least one edge from either colour i or colour j. Applying now Lemma 2.4.5, we deduce that with non-zero probability no event $A_{e,e'}$ occurs, as required. □

2.5.2 Properties of random subgraphs of the hypercube

In this section we state and prove some basic properties of random subgraphs of the hypercube. The first one guarantees that the degrees of all vertices are linear in the dimension.

Lemma 2.5.5. *Let $0 < \delta \ll \varepsilon \leq 1/2$. Then, we a.a.s. have that $\delta(\mathcal{Q}^n_{1/2+\varepsilon}) \geq \delta n$.*

Proof. Let $p \coloneqq 1/2 + \varepsilon$. Fix any $v \in \{0,1\}^n$. Throughout this proof, we write $d(v)$ to refer to the degree of v in $\mathcal{Q}^n_p$.

Note that $d(v)$ follows a binomial distribution with parameters n and p. Since $\delta < 1/2$, it follows that

$$\mathbb{P}[d(v) \leq \delta n] \leq \delta n \binom{n}{\delta n} p^{\delta n}(1-p)^{(1-\delta)n}.$$

Using the Stirling formula, we conclude that

$$\mathbb{P}[d(v) \leq \delta n] \leq (1 + \mathcal{O}(n^{-1}))\sqrt{\frac{\delta n}{2\pi(1-\delta)}} \left(\left(\frac{p}{\delta}\right)^{\delta} \left(\frac{1-p}{1-\delta}\right)^{1-\delta} \right)^{n}.$$

By the union bound, it now suffices to show that

$$\left(\frac{p}{\delta}\right)^{\delta} \left(\frac{1-p}{1-\delta}\right)^{1-\delta} = \left(\frac{1+2\varepsilon}{2\delta}\right)^{\delta} \left(\frac{1-2\varepsilon}{2(1-\delta)}\right)^{1-\delta} < \frac{1}{2},$$

but this follows since $\delta \ll \varepsilon$. $\square$

Next we show that, in any ball of radius ℓ, the number of vertices whose degree is far from the expected is much smaller (at most a constant) if we allow larger deviations for the degrees. Even more, we can prove a similar statement if we restrict the degrees to some linear subsets of the total neighbourhood in $\mathcal{Q}^n$. Recall that, for any vertex $v \in \{0,1\}^n$, any graph $G \subseteq \mathcal{Q}^n$ and a set $S \subseteq N_{\mathcal{Q}^n}(v)$, we denote $d_G(v, S) = |N_G(v) \cap S|$.

Lemma 2.5.6. *Let $\varepsilon, \delta, \gamma \in (0,1)$ and $\ell \in \mathbb{N}$. For each $v \in \{0,1\}^n$, let $S(v) \subseteq N_{\mathcal{Q}^n}(v)$ satisfy $|S(v)| \geq \gamma n$. Let $\mathcal{E}$ be the event that there are no vertices $v \in \{0,1\}^n$ for which $|\{u \in B^{\ell}(v) : d_{\mathcal{Q}^n_{\varepsilon}}(u, S(u)) \neq (1 \pm \delta)\varepsilon|S(u)|\}| \geq 100/(\delta^2\varepsilon\gamma)$. Then, for n sufficiently large, $\mathbb{P}[\mathcal{E}] \geq 1 - e^{-4n}$.*

Proof. Throughout this proof, we write $d(v)$ for $d_{\mathcal{Q}^n_{\varepsilon}}(v)$ and $d(v, S)$ for $d_{\mathcal{Q}^n_{\varepsilon}}(v, S)$, for any set S.

Let $C := \lceil 100/(\delta^2\varepsilon\gamma) \rceil$. Fix any vertex $v \in \{0,1\}^n$ and $A \in \binom{B^{\ell}(v)}{C}$. Observe that for any $u \in A$, if $d(u, S(u)) \neq (1 \pm \delta)\varepsilon|S(u)|$, then $d(u, S(u) \setminus A) \neq (1 \pm \delta/2)\varepsilon|S(u)|$. Observe that $\mathbb{E}[d(u, S(u) \setminus A)] \in [\varepsilon(|S(u)| - C), \varepsilon|S(u)|]$ for all $u \in A$. Furthermore, the variables $\{d(u, S(u) \setminus A) : u \in A\}$ are mutually independent, and each of them follows a binomial distribution. By Lemma 2.4.2, for each $u \in A$ we have that, for n sufficiently large,

$$\mathbb{P}[d(u) \neq (1 \pm \delta)\varepsilon|S(u)|] \leq \mathbb{P}[d(u, S(u) \setminus A) \neq (1 \pm \delta/2)\varepsilon|S(u)|] \leq 2e^{-\delta^2\varepsilon\gamma n/19} \leq e^{-\delta^2\varepsilon\gamma n/20}.$$

We say that A is bad if $d(u, S(u)) \neq (1 \pm \delta)\varepsilon|S(u)|$ for all $u \in A$. Since the variables $d(u, S(u) \setminus A)$ are mutually independent, it follows that

$$\mathbb{P}[A \text{ is bad}] \leq \left(e^{-\delta^2\varepsilon\gamma n/20}\right)^C \leq e^{-5n}.$$

Observe that $\mathcal{E}$ holds if there are no bad sets A. By a union bound over all choices of v and all choices of A, it follows that

$$\mathbb{P}[\overline{\mathcal{E}}] \leq 2^n \binom{\ell n^\ell}{C} e^{-5n} \leq e^{-4n}. \qquad \square$$

Finally, due to word constraints, we state without proof Theorem 2.5.7. A full proof can be found in [31]. Roughly speaking, Theorem 2.5.7 states that, for any constant $\varepsilon > 0$ and $\ell \in \mathbb{N}$, with high probability the random graph $\mathcal{Q}^n_\varepsilon$ contains a set of ℓ-dimensional cubes which are vertex-disjoint, cover all but a small proportion of the vertices of $\mathcal{Q}^n_\varepsilon$, and are 'sufficiently significant' with respect to every large set of directions, while not being 'too significant' with respect to any given direction.

Given any $\ell \in \mathbb{N}$, any $S \subseteq \mathcal{D}(\mathcal{Q}^n)$ and any copy C of $\mathcal{Q}^\ell$ with $C \subseteq \mathcal{Q}^n$, we define the *significance of* C *in* S as $\sigma(C, S) \coloneqq |\mathcal{D}(C) \cap S|$. Similarly, given any set $\mathcal{C}$ of ℓ-dimensional cubes in $\mathcal{Q}^n$, we define the *significance of* $\mathcal{C}$ *in* S as $\sigma(\mathcal{C}, S) \coloneqq \sum_{C \in \mathcal{C}} \sigma(C, S)$. We also denote $\Sigma(\mathcal{C}, S, t) \coloneqq \{C \in \mathcal{C} : \sigma(C, S) \geq t\}$. Given any $x \in \{0,1\}^n$ and any $Y \subseteq N_{\mathcal{Q}^n}(x)$, we denote $\mathcal{C}_x(Y) \coloneqq \{C \in \mathcal{C} : \mathrm{dist}(x, C) = 1, V(C) \cap Y \neq \varnothing\}$. In particular, we will write $\mathcal{C}_x \coloneqq \mathcal{C}_x(N_{\mathcal{Q}^n}(x))$.

Theorem 2.5.7. *Let $\varepsilon, \delta, \alpha, \beta \in (0,1)$ and $K, \ell \in \mathbb{N}$ be such that $1/\ell \ll \alpha \ll \beta$. For each $x \in \{0,1\}^n$, let $A_0(x) \coloneqq N_{\mathcal{Q}^n}(x)$ and, for each $i \in [K]$, let $A_i(x) \subseteq A_0(x)$ be a set of size $|A_i(x)| \geq \beta n$. Then, the graph $\mathcal{Q}^n_\varepsilon$ a.a.s. contains a collection $\mathcal{C}$ of vertex-disjoint copies of $\mathcal{Q}^\ell$ such that the following properties are satisfied for every $x \in \{0,1\}^n$:*

(M1) $|A_0(x) \cap V(\mathcal{C})| \geq (1-\delta)n$;

(M2) *for every* $\hat{e} \in \mathcal{D}(\mathcal{Q}^n)$ *we have* $|\Sigma(\mathcal{C}_x, \{\hat{e}\}, 1)| = o(n^{1/2})$;

(M3) *for every* $i \in [K]_0$ *and every* $S \subseteq \mathcal{D}(\mathcal{Q}^n)$ *with* $\alpha n/2 \leq |S| \leq \alpha n$ *we have*

$$|\Sigma(\mathcal{C}_x(A_i(x)), S, \ell^{1/2})| \geq |A_i(x)|/3000.$$

2.6 Near-spanning trees in random subgraphs of the hypercube

In this section we present our results on bounded degree near-spanning trees in $\mathcal{Q}^n_\varepsilon$. In Section 2.6.1 we prove the main result of this section (Theorem 2.6.1). This implies that with high probability there exists a near-spanning bounded degree tree in $\mathcal{Q}^n_\varepsilon$, which covers most of the neighbourhood of every vertex whilst avoiding a small random set of vertices, to which we refer as a reservoir. In Section 2.6.2 we prove Theorem 2.6.19, which allows us to extend the tree using vertices of the reservoir such that (amongst others) the proportion of uncovered vertices is even smaller. Finally, in Lemma 2.6.20 we show that, if some number of small local obstructions is prescribed, the tree given by Theorem 2.6.19 can be slightly modified to avoid these obstructions. For convenience, throughout this section, we move away from the algebraic notation for the hypercube to a more combinatorial notation.

We (re-)define the hypercube by setting $V(\mathcal{Q}^n) := \mathcal{P}([n])$ and joining two vertices $u, v \in \mathcal{P}([n])$ by an edge if and only if $||u| - |v|| = 1$ and $u \subseteq v$ or $v \subseteq u$. In this setting, directions correspond to the elements in $[n]$, and following a direction $i \in [n]$ from a vertex $v \in \mathcal{P}([n])$ means adding i to v if $i \notin v$, or deleting it from v if $i \in v$. Note that there is a natural partition of $V(\mathcal{Q}^n)$ into sets such that every vertex of a set has the same size. Given any set $S \subseteq [n]$, we denote $S^{(t)} := \{X \subseteq S : |X| = t\}$. We will denote by L_i, for $i \in [n]_0$, the set of all vertices $v \in V(\mathcal{Q}^n) = \mathcal{P}([n])$ with $|v| = i$ (that is, $L_i = [n]^{(i)}$), and we will refer to these sets as *levels*. This notation is especially useful because of the natural notion of containment of vertices, which provides a partial order on the vertices of $\mathcal{Q}^n$. Given any graph $G \subseteq \mathcal{Q}^n$, for a vertex $x \in L_i$, we refer to the neighbours of x in G lying in L_{i+1} as *up-neighbours*, and to the neighbours of x in L_{i-1} as *down-neighbours*, and denote these sets by $N_G^{\uparrow}(x)$ and $N_G^{\downarrow}(x)$, respectively. We write $d_G^{\uparrow}(x) \coloneqq |N_G^{\uparrow}(x)|$ and $d_G^{\downarrow}(x) \coloneqq |N_G^{\downarrow}(x)|$. Whenever the subscript is omitted, we mean that $G = \mathcal{Q}^n$. We will say that a path $P = v_1 \dots v_k$ in $\mathcal{Q}^n$ is a *chain* if its vertices satisfy the relation $v_1 \subsetneq \dots \subsetneq v_k$, and refer to it as a v_1-v_k chain.

In more generality, because of the symmetries of the hypercube, this notation can be extended with respect to any vertex $v \in V(\mathcal{Q}^n)$ by defining, for each $i \in [n]_0$, $L_i(v) \coloneqq \{u \in V(\mathcal{Q}^n) : \mathrm{dist}(u, v) = i\}$. One can then define up-neighbours and down-neighbours with respect to v, and use the notations $N_G^{\uparrow}(x, v)$, $N_G^{\downarrow}(x, v)$, $d_G^{\uparrow}(x, v)$ and $d_G^{\downarrow}(x, v)$, for $G \subseteq \mathcal{Q}^n$. We say that a path $P = v_1 \dots v_k$ in $\mathcal{Q}^n$ is a *chain with respect to* v if its vertices satisfy that, if $v_1 \in L_j(v)$ for some $j \in [n]_0$, then for all $\ell \in [k] \setminus \{1\}$ we have $v_\ell \in L_{j+\ell-1}(v)$, and refer to it as a v_1-v_k chain. Given any graph $G \subseteq \mathcal{Q}^n$, for any $i \in [n]$ and $v \in V(\mathcal{Q}^n)$, we will write $E_G(L_{i-1}(v), L_i(v))$ for the set of edges of G whose endpoints lie in the levels $L_{i-1}(v)$ and $L_i(v)$, respectively. We will drop the subscript whenever $G = \mathcal{Q}^n$.

2.6.1 Constructing a bounded degree near-spanning tree

Our goal in this subsection is to prove Theorem 2.6.1 below. Given a graph G and $\delta \in [0, 1]$, let $Res(G, \delta)$ be a probability distribution on subsets of $V(G)$, where $R \sim Res(G, \delta)$ is obtained by adding each vertex $v \in V(G)$ to R with probability δ, independently of every other vertex. We will refer to this set R as a *reservoir*.

Theorem 2.6.1. *Let $0 < 1/D, \delta \ll \varepsilon' \leq 1/2$, and let $\varepsilon, \gamma \in (0, 1]$ and $k \in \mathbb{N}$. Then, the following holds a.a.s. Let $\mathcal{S} \subseteq V(\mathcal{Q}^n)$ with the following two properties:*

(P1) *for any distinct $x, y \in \mathcal{S}$ we have $\mathrm{dist}(x, y) \geq \gamma n$, and*

(P2) *$B_{\mathcal{Q}^n}^{k+2}(\mathcal{S}) \cap \{\varnothing, [n], [\lceil n/2 \rceil], [n] \setminus [\lceil n/2 \rceil]\} = \varnothing$.*

Let $R \sim Res(\mathcal{Q}^n, \delta)$. Then, there exists a tree $T \subseteq \mathcal{Q}^n_\varepsilon - (R \cup B^k_{\mathcal{Q}^n}(\mathcal{S}))$ such that

(T1) $\Delta(T) < D$,

(T2) *for all $x \in V(\mathcal{Q}^n) \setminus B^k_{\mathcal{Q}^n}(\mathcal{S})$, we have that $|N_{\mathcal{Q}^n}(x) \cap V(T)| \geq (1 - \varepsilon')n$.*

The set $\mathcal{S}$ will be important in the proof of Theorems 2.1.5 and 2.1.8, where it will play the role of the set $\mathcal{U}$ of vertices of small degree. In the proof of Theorems 2.1.1 and 2.1.7 we can take $\mathcal{S} = \varnothing$.

To prove Theorem 2.6.1, we will consider suitable 'branching-like' processes which start at the 'bottom' of the hypercube, and grow 'upwards'. The tree will be formed by considering unions of such processes. The precise definition of the model we use is given in Definition 2.6.3. Crucially, there is a joint distribution of this branching-like process model and the binomial model $\mathcal{Q}^n_\varepsilon$. These processes are analysed and constructed in the results leading up to Lemma 2.6.12. Subgraphs obtained from the processes are then connected into a tree in Lemma 2.6.17.

We begin with a formal description of our model. We denote by $\mathbf{p} = (p_0, \dots, p_{n-1}) \in [0,1]^n$ an n-component vector of probabilities. We now describe a distribution on subgraphs of $\mathcal{Q}^n$ which is biased with respect to the number of edges between different levels of the hypercube.

Definition 2.6.2 (Level-biased subgraphs of $\mathcal{Q}^n$)**.** *Given $n \in \mathbb{N}$ and $\mathbf{p} = (p_0, \dots, p_{n-1}) \in [0,1]^n$, let $\mathcal{W}^n_{\mathbf{p}}$ be a distribution on subgraphs of $\mathcal{Q}^n$ where $W \sim \mathcal{W}^n_{\mathbf{p}}$ is generated as follows: we set $V(W) \coloneqq V(\mathcal{Q}^n)$ and, for each $i \in [n-1]_0$, each $e \in E(L_i, L_{i+1})$ is included in W with probability p_i, independently of all other edges.*

Roughly speaking, the above model has the advantage that, by choosing our probabilities p_i appropriately, it will allow us to generate subgraphs of $\mathcal{Q}^n$ where each vertex has the same number of up-neighbours in expectation. Moreover, note that there is a joint distribution of $\mathcal{W}^n_{\mathbf{p}}$ and $\mathcal{Q}^n_p$ such that we have $\mathcal{W}^n_{\mathbf{p}} \subseteq \mathcal{Q}^n_p$, where p is the maximum component of $\mathbf{p}$.

We are now in a position to define one further distribution on subgraphs of $\mathcal{Q}^n$. We will search for a near-spanning tree for $\mathcal{Q}^n_\varepsilon$ in the graphs generated according to this distribution.

Definition 2.6.3 (Percolation graph $\mathcal{P}(n, \mathbf{p}, M)$)**.** *Given $n, M \in \mathbb{N}$ and $\mathbf{p} = (p_0, \ldots, p_{n-1}) \in [0,1]^n$, we define $\mathcal{P}(n, \mathbf{p}, M)$ to be a distribution on subgraphs of $\mathcal{Q}^n$ where $P \sim \mathcal{P}(n, \mathbf{p}, M)$ is generated as follows. Let $R \sim Res(\mathcal{Q}^n, 1/100)$ and $W \sim \mathcal{W}^n_{\mathbf{p}}$. For each $x \in V(\mathcal{Q}^n)$, if $d^{\uparrow}_W(x) \geq M$, let $B(x) \subseteq N^{\uparrow}_W(x)$ be a uniformly random set of size M (otherwise, let $B(x) := \varnothing$), and let $E(x)$ be the set of edges joining x to each $y \in B(x)$. Let W' be the spanning subgraph of W with edge set $\bigcup_{x \in V(\mathcal{Q}^n)} E(x)$. The graph $P \subseteq \mathcal{Q}^n$ is then given by setting $P := W' - R$.*

Remark 2.6.4. *Observe that, given any two distinct edges $e, e' \in E(\mathcal{Q}^n)$, the events $e \in E(W')$ and $e' \in E(W')$ are mutually dependent if and only if for some $i \in [n]$ we have $e, e' \in E(L_{i-1}, L_i)$ with $e \cap e' = \{v\}$ for some $v \in L_{i-1}$. Otherwise, these events are independent. In particular, if $e \in E(L_{i-1}, L_i)$ and $e' \in E(L_{j-1}, L_j)$ with $i \neq j$, then these events are always independent.*

Note that $\mathcal{P}(n, \mathbf{p}, M) \subseteq \mathcal{W}^n_{\mathbf{p}}$ by definition, and therefore we have a joint distribution of $\mathcal{P}(n, \mathbf{p}, M)$ and $\mathcal{Q}^n_p$ such that $\mathcal{P}(n, \mathbf{p}, M) \subseteq \mathcal{Q}^n_p$, where p is the maximum component of $\mathbf{p}$.

Definition 2.6.5 (Feasible $(n, \mathbf{p}, M)$)**.** *We say that the tuple $(n, \mathbf{p}, M)$ is* feasible *if*

(i) *$p_i = 0$ for all $9n/10 < i < n$,*

(ii) *$max_{i \in [n-1]_0} p_i < 1/10$ and $M > 1600$,*

(iii) *there exists $t \in \mathbb{R}$ with $600 < t < 100M$ such that $P \sim \mathcal{P}(n, \mathbf{p}, M)$ satisfies $\mathbb{P}[e \in E(P)] = t/n$ for all $e \in \bigcup_{i=0}^{\lfloor 9n/10 \rfloor} E(L_i, L_{i+1})$.*

Remark 2.6.6. *Let $(n, \mathbf{p}, M)$ be feasible, where $\mathbf{p} = (p_0, \ldots, p_{n-1})$. Note that p_0 determines the value of p_i for all $i \in [\lfloor 9n/10 \rfloor]$. Furthermore, let $P \sim \mathcal{P}(n, \mathbf{p}, M)$. We can generate P by first sampling $W \sim \mathcal{W}^n_{\mathbf{p}}$ and $R \sim Res(\mathcal{Q}^n, 1/100)$, and then defining the graph W' as* *described in Definition 2.6.3. Let* $t' := t/(\frac{99}{100})^2$*, where* t *is as in Definition 2.6.5(iii). Since* $(n, \mathbf{p}, M)$ *is feasible, for all* $e \in \bigcup_{i=0}^{\lfloor 9n/10 \rfloor} E(L_i, L_{i+1})$ *we have*

$$\mathbb{P}[e \in E(W')] = t'/n.$$

Furthermore, for all $i \in [\lfloor 9n/10 \rfloor]$*, given* $e, e' \in E(L_i, L_{i+1})$ *with* $e \neq e'$*, we have that*

$$\mathbb{P}[e, e' \in E(W')] \leq \mathbb{P}[e \in E(W')]^2 = (t'/n)^2.$$

From here on, where it is clear from the context, we will use $p_0, \ldots, p_{n-1}$ to denote the components of each probability vector $\mathbf{p}$, and will use t to denote the value t in Definition 2.6.5 and t' to denote the value t' in Remark 2.6.6.

Proposition 2.6.7. *For all* $\varepsilon \in (0, 1/10)$*,* $M > 1600$*, and* $n \in \mathbb{N}$ *such that* $0 < 1/n \ll 1/M, \varepsilon$ *there exists a tuple* $(n, \mathbf{p}, M)$ *which is feasible and such that* $p_i \leq \varepsilon$ *for all* $i \in [n-1]_0$*.*

Proof. Let $P \sim \mathcal{P}(n, \mathbf{p}, M)$, for some $\mathbf{p}$ which will be determined later. We generate P by first sampling $W \sim \mathcal{W}_{\mathbf{p}}^n$ and $R \sim \mathit{Res}(\mathcal{Q}^n, 1/100)$. Let $j \in [\lfloor 9n/10 \rfloor]_0$ be fixed and let $e \in E(L_j, L_{j+1})$. Let $x \in L_j$ be incident with e. Let $\mathcal{A}$ be the event that $e \in E(W)$. For each $k \in [n-j]_0$, let $\mathcal{B}_k$ be the event that $d_W^{\uparrow}(x) = k$. Let $\mathcal{C}$ be the event that $e \in E(P)$. For each $i \in [n-M]_0$, let

$$f_i(y) := \left(\frac{99}{100}\right)^2 \frac{M}{n-i} \sum_{k=M}^{n-i} \binom{n-i}{k} y^k (1-y)^{n-i-k}.$$

Then, we have that

$$\begin{aligned}\mathbb{P}[\mathcal{C}] &= \sum_{k=M}^{n-j} \mathbb{P}[\mathcal{C} \mid \mathcal{A} \wedge \mathcal{B}_k]\, \mathbb{P}[\mathcal{A} \mid \mathcal{B}_k]\, \mathbb{P}[\mathcal{B}_k] \\ &= \sum_{k=M}^{n-j} \left(\frac{99}{100}\right)^2 \frac{M}{k} \frac{k}{n-j} \binom{n-j}{k} p_j^k (1-p_j)^{n-j-k} = f_j(p_j).\end{aligned}$$

Let $m := \min_{i \in [\lfloor 9n/10 \rfloor]_0} f_i(\varepsilon)$.

Claim 2.6.1. *We have* $\frac{600}{n} < m < \frac{100M}{n}$*.*

Proof of Claim 2.6.1. Let $i \in [\lfloor 9n/10 \rfloor]_0$ be such that $f_i(\varepsilon) = m$. Clearly,

$$f_i(\varepsilon) \leq \frac{100M}{n} \sum_{k=M}^{n-i} \binom{n-i}{k} \varepsilon^k (1-\varepsilon)^{n-i-k} < \frac{100M}{n}.$$

Moreover, we have that

$$f_i(\varepsilon) = \left(\frac{99}{100}\right)^2 \frac{M}{n-i} \sum_{k=M}^{n-i} \binom{n-i}{k} \varepsilon^k (1-\varepsilon)^{n-i-k} > \frac{1}{2} \left(\frac{99}{100}\right)^2 \frac{M}{n},$$

as $M \leq \varepsilon(n-i)/2$. ◀

For each $i \in [\lfloor 9n/10 \rfloor]_0$ such that $f_i(\varepsilon) > m$, since $f_i(x)$ is continuous, by the intermediate value theorem we can choose some $p_i \in (0, \varepsilon)$ such that $f_i(p_i) = m$. This determines the probability vector $\mathbf{p} = (p_0, \ldots, p_{n-1})$. By Claim 2.6.1, the tuple $(n, \mathbf{p}, M)$ is feasible (with mn playing the role of t in Definition 2.6.5), hence the statement is satisfied. □

In order to construct the near-spanning tree, we will generate a graph $P \sim \mathcal{P}(n, \mathbf{p}, M)$, for some feasible $(n, \mathbf{p}, M)$, and will be interested in whether or not there exists a chain in P from some vertex $x \in L_m$ to some vertex $y \in L_{m'}$, for $m' > m$ and $x \subseteq y$. Note that the presence and absence of such chains in P are highly dependent. Thus, in order to show that such chains exist with high probability, we will consider the number of x-y chains and bound its variance. We do so in the following lemma. In order to state it, we first need to set up some notation.

Given $x \in L_m$ and $y \in L_{m'}$ with $m' \geq m$, we denote by $\mathcal{X}_{x,y}$ the collection of x-y chains in $\mathcal{Q}^n$. For each $X \in \mathcal{X}_{x,y}$ and any graph $G \subseteq \mathcal{Q}^n$, let $Y_X(G)$ be the corresponding indicator variable which takes value 1 if $X \subseteq G$ and 0 otherwise. Let $Y_{x,y}(G) \coloneqq \sum_{X \in \mathcal{X}_{x,y}} Y_X(G)$. Whenever G is clear from the context, we will simply write $Y_{x,y}$. We define

$$\Delta(Y_{x,y}) \coloneqq \sum_{\substack{(X,X') \in \mathcal{X}_{x,y}^2 \\ X \neq X'}} \mathrm{Cov}[Y_X Y_{X'}],$$

so $\mathrm{Var}[Y_{x,y}] = \Delta(Y_{x,y}) + \sum_{X \in \mathcal{X}_{x,y}} \mathrm{Var}[Y_X]$.

Lemma 2.6.8. *Let $P \sim \mathcal{P}(n, \mathbf{p}, M)$, where $(n, \mathbf{p}, M)$ is feasible with $0 < 1/n \ll 1/M$. Let $1 \leq m < m' \leq 9n/10$ with $m' - m + 1 \geq n/4 - 1$. Let $x \in L_m$ and $y \in L_{m'}$ with $x \subseteq y$. Then,*

$$\Delta(Y_{x,y}) \leq 2\mathbb{E}[Y_{x,y}]^2.$$

The proof of Lemma 2.6.8 makes use of the analysis in the proof of a similar lemma of Kohayakawa, Kreuter, and Osthus [66, Lemma 7]. In order to shorten our analysis here, we first state a partial result which follows from the analysis of [66]. For this, we first need to give some more definitions.

Fix $x \in L_m$ and $y \in L_{m'}$ with $m' \geq m$ and $x \subseteq y$. Observe that $|\mathcal{X}_{x,y}| = \text{dist}(x, y)! = (m' - m)!$ depends only on the distance between x and y. For each $k \in [n]$, let $R_k := (k-1)!$. Given any $X, X' \in \mathcal{X}_{x,y}$ with $X \neq X'$, let $i(X, X') := |V(X) \cap V(X')| - 2$, let $s(X, X')$ be the number of connected components of $X - V(X')$, and let $\ell(X, X')$ be the largest order over these components.

Next, we define the set of possible intersection patterns for two chains. Let $k := m' - m + 1$. Given any chains $X, X' \in \mathcal{X}_{x,y}$, let $A(X, X')$ be the collection of indices $a \in [k-2]$ for which X and X' agree on their $(a+1)$-th elements (where we consider x to be the first element of X and X'). An *admissible* (i, ℓ, s)*-pattern* is a set $A \subseteq [k-2]$ with $|A| = i$ such that the longest interval of consecutive elements in $[k-2] \setminus A$ contains exactly ℓ elements and such that the number of maximal intervals of consecutive elements in $[k-2] \setminus A$ is exactly s. We denote by $\mathcal{A}_{i,\ell,s}$ the set of all admissible (i, ℓ, s)-patterns. Furthermore, we define $C_{i,\ell,s} := |\mathcal{A}_{i,\ell,s}|$. Note that any pair of chains $X, X' \in \mathcal{X}_{x,y}$ with $i(X, X') = i$, $\ell(X, X') = \ell$ and $s(X, X') = s$ define an admissible (i, ℓ, s)-pattern $A(X, X') \in \mathcal{A}_{i,\ell,s}$.

Given a chain $X \in \mathcal{X}_{x,y}$ and a pattern $A \in \mathcal{A}_{i,\ell,s}$, let $F(A)$ be the number of chains $X' \in \mathcal{X}_{x,y}$ such that $A(X, X') = A$. (Note that the definition of $F(A)$ is independent of X.) Let $F_{i,\ell,s} := \max_{A \in \mathcal{A}_{i,\ell,s}} F(A)$. Observe that $F_{i,\ell,s}$ is an upper bound on the number of chains X' with $A(X, X') = A$.

Finally, for each triple $(i, \ell, s) \in [k-3]_0 \times [k-2]^2$, let

$$\Delta_{i,\ell,s} := \sum_{\substack{(X,X')\in\mathcal{X}^2_{x,y},\, X\neq X' \\ i(X,X')=i,\, \ell(X,X')=\ell,\, s(X,X')=s}} \mathbb{E}[Y_X Y_{X'}].$$

Furthermore, let

$$\Delta_0(Y_{x,y}) := \sum_{\substack{(X,X')\in\mathcal{X}^2_{x,y} \\ i(X,X')=0}} \mathrm{Cov}[Y_X Y_{X'}] \qquad \text{and} \qquad \Delta_1(Y_{x,y}) := \sum_{\substack{(X,X')\in\mathcal{X}^2_{x,y} \\ i(X,X')\in[k-3]}} \mathrm{Cov}[Y_X Y_{X'}].$$

Thus, $\Delta(Y_{x,y}) = \Delta_0(Y_{x,y}) + \Delta_1(Y_{x,y})$. Note that, by summing $\Delta_{i,\ell,s}$ over all triples $(i, \ell, s) \in [k-3] \times [k-2]^2$, we obtain an upper bound for $\Delta_1(Y_{x,y})$.

Lemma 2.6.9 ([66])**.** *For all $M > 100$ there exists n_0 such that, for all $n \geq n_0$, the following holds. Let $x \in L_1$ and $y \in L_{n-1}$ with $x \subseteq y$. Let $p \geq M/(2n)$. Let $Q \subseteq \mathcal{Q}^n$ be a random subgraph chosen according to any distribution such that*

$$\frac{\Delta_{i,\ell,s}}{\mathbb{E}[Y_{x,y}]^2} \leq \frac{C_{i,\ell,s} F_{i,\ell,s}}{R_{n-1} p^i},$$

for each possible choice of $(i, \ell, s) \in [k-3] \times [k-2]^2$. Then,

$$\Delta_1(Y_{x,y}) \leq \frac{100}{M} \mathbb{E}[Y_{x,y}]^2.$$

With this, we are finally ready to prove Lemma 2.6.8.

Proof of Lemma 2.6.8. Let $P \sim \mathcal{P}(n, \mathbf{p}, M)$, where $(n, \mathbf{p}, M)$ is feasible. Recall, from Definition 2.6.3, that P is generated by first sampling a set $R \sim \mathit{Res}(\mathcal{Q}^n, 1/100)$ and a graph $W \sim \mathcal{W}^n_{\mathbf{p}}$. We then generate the graph W' by choosing, for each $v \in \bigcup_{i=0}^{\lfloor 9n/10 \rfloor} L_i$, a set of M up-neighbours uniformly at random from the set of up-neighbours v has in W, provided $d^{\uparrow}_{W'}(v) \geq M$ (and by setting $d^{\uparrow}_{W'}(v) := 0$ otherwise). Let $t' := t/(\frac{99}{100})^2$. Thus, for all $e \in \bigcup_{i=0}^{\lfloor 9n/10 \rfloor} E(L_i, L_{i+1})$ we have by Remark 2.6.6 that

$$\mathbb{P}[e \in W'] = t'/n.$$

Let $k \coloneqq m' - m + 1$, and let X be a fixed x-y chain in $\mathcal{Q}^n$. By Remark 2.6.4 it follows that

$$\mathbb{E}[Y_{x,y}] = R_k \mathbb{P}[X \subseteq P] = R_k (t'/n)^{k-1} \Big(\frac{99}{100}\Big)^k. \tag{2.6.1}$$

Furthermore, for all $(i, \ell, s) \in [k-3]_0 \times [k-2]^2$, we have that

$$\Delta_{i,\ell,s} \leq R_k C_{i,\ell,s} F_{i,\ell,s} \left(\frac{99t'}{100n}\right)^{2k-i-2}. \tag{2.6.2}$$

To see this, note that we may first choose an x-y chain X, for which there are R_k choices. Next, we choose an admissible (i, ℓ, s)-pattern $A \in \mathcal{A}_{i,\ell,s}$, of which there are $C_{i,\ell,s}$. We then have at most $F_{i,\ell,s}$ choices for x-y chains X' with $A(X, X') = A$. Next, we bound the number of vertices and edges of $X \cup X'$. It is clear that X has k vertices and $k-1$ edges, and $|V(X') \setminus V(X)| = k - i - 2$. Moreover, observe that $|E(X') \setminus E(X)| \geq k - i - 1$. The bound finally follows by considering the probability that all these vertices and edges are present in P and by Remark 2.6.6.

We are going to compute bounds for $\Delta_0(Y_{x,y})$ and $\Delta_1(Y_{x,y})$ separately, and then combine them to obtain the result. We begin with a bound for $\Delta_1(Y_{x,y})$. Combining (2.6.1) and (2.6.2), it follows that, for all $(i, \ell, s) \in [k-3] \times [k-2]^2$,

$$\frac{\Delta_{i,\ell,s}}{\mathbb{E}[Y_{x,y}]^2} \leq \frac{C_{i,\ell,s} F_{i,\ell,s}}{R_k} \left(\frac{n}{t'}\right)^i \left(\frac{1}{\frac{99}{100}}\right)^{i+2} \leq \frac{C_{i,\ell,s} F_{i,\ell,s}}{R_k} \left(\frac{n}{t'(\frac{99}{100})^3}\right)^i.$$

Note that $(\frac{m'-m+2}{n}) t' (\frac{99}{100})^3 > 100$. It follows that we can apply Lemma 2.6.9 with $(\frac{m'-m+2}{n}) t' (\frac{99}{100})^3$ and $m' - m + 2$ playing the roles of M and n and $p = t'(\frac{99}{100})^3 / n$ to obtain that

$$\Delta_1(Y_{x,y}) \leq \frac{100}{(\frac{m'-m+2}{n}) t' (\frac{99}{100})^3} \mathbb{E}[Y_{x,y}]^2 \leq \mathbb{E}[Y_{x,y}]^2. \tag{2.6.3}$$

We now turn our attention to $\Delta_0(Y_{x,y})$. For any two chains $X, X' \in \mathcal{X}_{x,y}$ such that $i(X, X') = 0$, we have that $X \cup X'$ has $2k - 2$ vertices and the same number of edges. Therefore, by Remarks 2.6.4 and 2.6.6 we have $\mathbb{E}[Y_X Y_{X'}] \leq (\frac{99t'}{100n})^{2k-2}$, and by (2.6.1) we have that

$$\Delta_0(Y_{x,y}) \leq R_k^2 \Big(\frac{99t'}{100n}\Big)^{2k-2} \Big(1 - \Big(\frac{99}{100}\Big)^2\Big) \leq \mathbb{E}[Y_{x,y}]^2. \tag{2.6.4}$$

The conclusion follows immediately by combining (2.6.3) and (2.6.4). □

In order to proceed further, we will consider unions of independent graphs $P \sim \mathcal{P}(n, \mathbf{p}, M)$.

Definition 2.6.10. *Let $n, M, C \in \mathbb{N}$ and $\mathbf{p} \in [0,1]^n$. We define $\mathcal{P}^C(n, \mathbf{p}, M)$ to be a distribution on subgraphs of $\mathcal{Q}^n$ such that $P \sim \mathcal{P}^C(n, \mathbf{p}, M)$ is generated by taking C independently generated graphs $P_i \sim \mathcal{P}(n, \mathbf{p}, M)$ and setting $P := \bigcup_{i=1}^C P_i$. For each $i \in [C]$, there is a set $R_i \sim Res(\mathcal{Q}^n, 1/100)$ associated with P_i. Let $R := \bigcap_{i=1}^C R_i$. We say that R is the* reservoir *associated with P.*

It follows from Definitions 2.6.3 and 2.6.10 that there is a joint distribution of $\mathcal{P}^C(n, \mathbf{p}, M)$ and $\mathcal{Q}^n_{\min\{1,Cp\}}$ such that $\mathcal{P}^C(n, \mathbf{p}, M) \subseteq \mathcal{Q}^n_{\min\{1,Cp\}}$, where $p = \max_{i \in [n-1]_0} p_i$. Note that for all $x \in V(\mathcal{Q}^n)$ we have that $\mathbb{P}[x \in R] = (1/100)^C$.

Our next goal is to prove that, by choosing constants appropriately, there is a high probability that there exists an x-y chain in $P \sim \mathcal{P}^C(n, \mathbf{p}, M)$, even if we restrict the set of 'valid' chains to a significant subset of the total. For this, we will make use of Lemma 2.6.8. Given any vertices $x \in L_m$ and $y \in L_{m'}$ with $x \subseteq y$, any set $\mathcal{Z} \subseteq \mathcal{X}_{x,y}$, and any graph $G \subseteq \mathcal{Q}^n$, we denote the number of x-y chains $X \in \mathcal{Z}$ such that $X \subseteq G$ by $Y(\mathcal{Z}, G)$.

Lemma 2.6.11. *For $n, C \in \mathbb{N}$ and $\eta, \alpha > 0$ such that $0 < 1/n \ll 1/C \ll \eta, \alpha$ and any feasible $(n, \mathbf{p}, M)$ with $0 < 1/n \ll 1/M$, the following holds. Let $1 \leq m < m' \leq 9n/10$ with $m' - m + 1 \geq n/4 - 1$. Let $x \in L_m$ and $y \in L_{m'}$ with $x \subseteq y$. Let $\mathcal{Z}_{x,y} \subseteq \mathcal{X}_{x,y}$ be such that $|\mathcal{Z}_{x,y}| \geq \alpha |\mathcal{X}_{x,y}|$. Let $P \sim \mathcal{P}^C(n, \mathbf{p}, M)$. Then,*

$$\mathbb{P}[Y(\mathcal{Z}_{x,y}, P) > 0] \geq 1 - \eta.$$

Proof. For each $i \in [C]$, let $P_i \sim \mathcal{P}(n, \mathbf{p}, M)$, and let $P := \bigcup_{i=1}^C P_i$. Let $Y_i := Y_{x,y}(P_i)$ and $Z_i := Y(\mathcal{Z}_{x,y}, P_i)$, and let

$$\Delta(Z_i) := \sum_{\substack{(X,X') \in \mathcal{Z}_{x,y}^2 \\ X \neq X'}} \mathrm{Cov}[Y_X(P_i) Y_{X'}(P_i)].$$

Note that

$$\mathbb{E}[Y_i^2] \le \Delta(Y_i) + \mathbb{E}[Y_i] + \mathbb{E}[Y_i]^2. \tag{2.6.5}$$

We also have

$$\mathbb{E}[Z_i^2] \le \mathbb{E}[Y_i^2] \tag{2.6.6}$$

and, since all x-y chains are equiprobable,

$$\mathbb{E}[Z_i]^2 \ge \alpha^2 \mathbb{E}[Y_i]^2. \tag{2.6.7}$$

Let $k := m' - m + 1$. By (2.6.1), we have that $\mathbb{E}[Y_i] = R_k (t'/n)^{k-1}(99/100)^k$, where t' is the value given in Remark 2.6.6. Recall that $R_k = |\mathcal{X}_{x,y}| = (k-1)!$. We have by Stirling's formula that $\mathbb{E}[Y_i] > 1$. Therefore, $\mathbb{E}[Y_i] \le \mathbb{E}[Y_i]^2$. Moreover, it follows by Lemma 2.6.8 that $\Delta(Y_i) \le 2\mathbb{E}[Y_i]^2$. So $\mathbb{E}[Y_i^2] \le 4\mathbb{E}[Y_i]^2$ by (2.6.5). Combining this with (2.6.6), (2.6.7) and Proposition 2.4.1 we obtain

$$\mathbb{P}[Z_i = 0] \le 1 - \frac{\mathbb{E}[Z_i]^2}{\mathbb{E}[Z_i^2]} \le 1 - \frac{\alpha^2 \mathbb{E}[Y_i]^2}{\mathbb{E}[Z_i^2]} \le 1 - \frac{\alpha^2 \mathbb{E}[Y_i]^2}{\mathbb{E}[Y_i^2]} \le 1 - \alpha^2/4.$$

It follows that

$$\mathbb{P}[Y(\mathcal{Z}_{x,y}, P) = 0] = \prod_{i \in [C]} \mathbb{P}[Z_i = 0] \le (1 - \alpha^2/4)^C \le \eta. \qquad \square$$

When performing our analysis on the structure of P, the dependence of chains on each other becomes difficult to take into account. In order to deal with this issue, we will show that, with high probability, it suffices to consider only chains which lie in some large subsets of the total sets of chains, with the property that the presence or absence of a chain in one of these large subsets is independent from chains of all other subsets. (Note that Lemma 2.6.11 works for these sets of chains as long as they are not too small.) Lemmas 2.6.12 and 2.6.16 guarantee the existence of such sets. In Lemma 2.6.12 we prove that, assuming $x, x' \in L_m$, and $y, y' \in L_{m'}$, where y, y' are far apart, one can construct very large sets of chains between the pairs x, y and x', y', which are independent in the sense described above. Then, in Lemma 2.6.16 we will prove that we can pick many endpoints $y \in L_{m'}$ in such a way that they are suitably far apart. Moreover, we will combine Proposition 2.6.14 with Lemma 2.6.15 to later show that these large sets of chains can be constructed in such a way that each avoids a small fixed set of forbidden vertices. This will be important in showing that the tree we construct for our main result can be made to avoid the forbidden set $\mathcal{S}$ in the statement of Theorem 2.6.1.

Given $0 \leq m < m' \leq n$, let $x, x' \in L_m$ and $y, y' \in L_{m'}$ with $x \subseteq y$ and $x' \subseteq y'$. We denote by $\mathcal{X}_{x,y}^{\neg x',y'}$ the collection of chains $X \in \mathcal{X}_{x,y}$ for which there is no $X' \in \mathcal{X}_{x',y'}$ with $V(X) \cap V(X') \neq \varnothing$.

Lemma 2.6.12. *For all $n \geq 100$, the following holds. Let $1 \leq m < m' \leq n-1$ be such that $n/4 - 1 \leq k := m' - m + 1 \leq n/2$. Let $x, x' \in L_m$ and $y, y' \in L_{m'}$ with $x \subseteq y$ and $x' \subseteq y'$ be such that $\operatorname{dist}(x, x') = 2$ and $\operatorname{dist}(y, y') \geq 9k^2/(10n)$. Then,*

$$|\mathcal{X}_{x,y}^{\neg x',y'}| \geq \left(1 - \frac{60000}{n}\right) |\mathcal{X}_{x,y}|.$$

Proof. We may assume that $x \cup x' \subseteq y \cap y'$, since otherwise $\mathcal{X}_{x,y}^{\neg x',y'} = \mathcal{X}_{x,y}$. Let $b := |y \cap y'|$. We have that $b \leq m' - 9k^2/(20n)$. Let H denote the smallest subcube of $\mathcal{Q}^n$ which contains both $x \cup x'$ and $y \cap y'$. For each $i \in [b] \setminus [m]$, let $\mathcal{X}_{x,y}^i \subseteq \mathcal{X}_{x,y}$ be the set of chains $X \in \mathcal{X}_{x,y}$ such that $V(X) \cap L_i \cap V(H) \neq \varnothing$. Note that $\mathcal{X}_{x,y}^{\neg x',y'} \supseteq \mathcal{X}_{x,y} \setminus \bigcup_{i \in [b] \setminus [m]} \mathcal{X}_{x,y}^i$ and

$$|\mathcal{X}_{x,y}^i| = \binom{b-m-1}{i-m-1} (m'-i)!(i-m)!. \tag{2.6.8}$$

Indeed, there are $\binom{b-m-1}{i-m-1}$ choices to fix an element $z \in V(H) \cap L_i$. (To see this, consider that H is itself a cube of dimension $b - m - 1$, and we are choosing a vertex z from the $(i-m-1)$-th level of this cube.) Then, there are $(i-m)!$ x-z chains, and $(m'-i)!$ z-y chains.

Recall that $|\mathcal{X}_{x,y}| = (k-1)!$. By comparing this with (2.6.8) and simplifying, for all $i \in [b] \setminus [m]$ we obtain

$$\frac{|\mathcal{X}_{x,y}^i|}{|\mathcal{X}_{x,y}|} = \frac{i-m}{m'-m} \prod_{j=1}^{i-m-1} \frac{b-m-j}{m'-m-j} \leq \frac{i-m}{k-1}. \tag{2.6.9}$$

We now split the analysis into two cases. First, when i is small, we bound (2.6.9) directly. For all $i \in [b] \setminus [m]$ with $i \leq m + 64$, it follows from (2.6.9) that

$$\frac{|\mathcal{X}_{x,y}^i|}{|\mathcal{X}_{x,y}|} \leq \frac{64}{n/4 - 1} \leq \frac{300}{n}. \tag{2.6.10}$$

On the other hand, for each $i \in [b-1] \setminus [m]$, by (2.6.9) we have that

$$\frac{|\mathcal{X}_{x,y}^{i+1}|}{|\mathcal{X}_{x,y}|} = \frac{|\mathcal{X}_{x,y}^i|}{|\mathcal{X}_{x,y}|} \frac{i+1-m}{i-m} \cdot \frac{b-i}{m'-i} \leq \frac{39}{40} \frac{i+1-m}{i-m} \frac{|\mathcal{X}_{x,y}^i|}{|\mathcal{X}_{x,y}|}.$$

For all $i \in [b-1] \setminus [m+64]$, this yields

$$\frac{|\mathcal{X}_{x,y}^{i+1}|}{|\mathcal{X}_{x,y}|} \leq \frac{99}{100} \frac{|\mathcal{X}_{x,y}^i|}{|\mathcal{X}_{x,y}|}. \tag{2.6.11}$$

Finally, by combining (2.6.10) and (2.6.11), and considering a geometric series we conclude that

$$\frac{|\mathcal{X}_{x,y}^{\neg x',y'}|}{|\mathcal{X}_{x,y}|} \geq \frac{|\mathcal{X}_{x,y}| - \sum_{i=m+1}^{b} |\mathcal{X}_{x,y}^i|}{|\mathcal{X}_{x,y}|} \geq 1 - \frac{60000}{n}. \qquad \square$$

Remark 2.6.13. *Lemma 2.6.12 holds similarly if* $\mathrm{dist}(x, x') \geq 9k^2/(10n)$ *and* $\mathrm{dist}(y, y') = 2$.

Proposition 2.6.14. *Let* $0 < 1/n \ll \gamma, 1/k \leq 1$*, where* $n, k \in \mathbb{N}$*, and let* $\mathcal{S} \subseteq V(\mathcal{Q}^n)$ *be such that, for all distinct* $x, x' \in \mathcal{S}$*, we have* $\mathrm{dist}(x, x') \geq \gamma n$*. Then, for any* $y \in L_m$ *such that* $m \geq n/8$*, and for every* $\gamma m/2 \leq t \leq (1 - \gamma/2)m$*, we have* $|y^{(t)} \cap B_{\mathcal{Q}^n}^k(\mathcal{S})| \leq |y^{(t)}| 2^{-\gamma n/200}$.

Proof. Let $m \geq n/8$ and $y \in L_m$. Let $\gamma m/2 \leq t \leq (1 - \gamma/2)m$ and let $\mathcal{S}' \subseteq \mathcal{S}$ be the set of all those $x \in \mathcal{S}$ for which $B_{\mathcal{Q}^n}^k(x) \cap y^{(t)} \neq \varnothing$. We have that

$$|B_{\mathcal{Q}^n}^k(\mathcal{S}) \cap y^{(t)}| = \sum_{x \in \mathcal{S}'} |B_{\mathcal{Q}^n}^k(x) \cap y^{(t)}| \leq 2n^k |\mathcal{S}'|. \tag{2.6.12}$$

Moreover, for every $x, x' \in \mathcal{S}'$ we have that $B_{\mathcal{Q}^n}^{\gamma n/3}(x) \cap B_{\mathcal{Q}^n}^{\gamma n/3}(x') = \varnothing$, and, therefore,

$$|\mathcal{S}'| (\min_{x \in \mathcal{S}'} |B_{\mathcal{Q}^n}^{\gamma n/3}(x) \cap y^{(t)}|) \leq |y^{(t)}|. \tag{2.6.13}$$

Claim 2.6.2. *For every* $x \in \mathcal{S}'$ *we have* $|B_{\mathcal{Q}^n}^{\gamma n/3}(x) \cap y^{(t)}| \geq 2^{\gamma m/20}$.

Proof. Let $x' \in B^k_{\mathcal{Q}^n}(x) \cap y^{(t)}$. Let $z \subseteq x'$ be such that $z \in L_{\gamma m/7}$ (recall that $t \geq \gamma m/2$). Since $y \in L_m$ we have that $|y \setminus x'| = m - t$. Let $z' \subseteq y \setminus x'$ be such that $z' \in L_{\gamma m/7}$ (recall that $t \leq (1-\gamma/2)m$). It follows that $(x' \setminus z) \cup z' \in B^{\gamma n/3}_{\mathcal{Q}^n}(x) \cap y^{(t)}$. Note that there are $\binom{t}{\gamma m/7}$ choices for z and $\binom{m-t}{\gamma m/7}$ choices for z'. It follows that

$$\left|B^{\gamma n/3}_{\mathcal{Q}^n}(x) \cap y^{(t)}\right| \geq \binom{m-t}{\gamma m/7}\binom{t}{\gamma m/7} \geq 2^{\gamma m/20}. \qquad \blacktriangleleft$$

Combining (2.6.12), (2.6.13) and the above claim we have

$$|B^k_{\mathcal{Q}^n}(\mathcal{S}) \cap y^{(t)}| \leq \frac{2n^k|y^{(t)}|}{\min_{x\in\mathcal{S}'}|B^{\gamma n/3}_{\mathcal{Q}^n}(x) \cap y^{(t)}|} \leq 2n^k|y^{(t)}|2^{-\gamma m/20} \leq |y^{(t)}|2^{-\gamma n/200}. \qquad \square$$

Given $x, y \in V(\mathcal{Q}^n)$ with $x \subseteq y$ and $\mathcal{S} \subseteq V(\mathcal{Q}^n)$, we denote by $X^{\neg\mathcal{S}}_{x,y}$ the collection of chains $X \in \mathcal{X}_{x,y}$ for which $V(X) \cap \mathcal{S} = \varnothing$.

Lemma 2.6.15. *Let $0 < 1/n \ll \gamma, 1/k \leq 1$ where $n, k \in \mathbb{N}$, and let $\mathcal{S} \subseteq V(\mathcal{Q}^n)$ be such that for all $x, x' \in \mathcal{S}$ we have $\mathrm{dist}(x, x') \geq \gamma n$. Let $x, y \in V(\mathcal{Q}^n) \setminus B^k_{\mathcal{Q}^n}(\mathcal{S})$ with $x \subseteq y$ and $m := \mathrm{dist}(x, y) \geq n/8$. Then, $|\mathcal{X}^{\neg B^k_{\mathcal{Q}^n}(\mathcal{S})}_{x,y}| \geq 3m!/4$.*

Proof. We may assume that $x = \varnothing$ and $y = [m]$, where $m \geq n/8$. Let $\mathcal{X}^i_{x,y}$ denote the collection of chains $X \in \mathcal{X}_{x.y}$ for which $V(X) \cap L_i \cap B^k_{\mathcal{Q}^n}(\mathcal{S}) \neq \varnothing$. We have

$$|\mathcal{X}_{x,y} \setminus \mathcal{X}^{\neg B^k_{\mathcal{Q}^n}(\mathcal{S})}_{x,y}| \leq \sum_{i=1}^{m-1} |\mathcal{X}^i_{x,y}|. \tag{2.6.14}$$

Furthermore, by Proposition 2.6.14 (with $\gamma/2$ playing the role of γ), for all $\gamma m/4 \leq i \leq (1-\gamma/4)m$ we have that

$$|\mathcal{X}^i_{x,y}| \leq \binom{m}{i} 2^{-\gamma m/400} i!(m-i)! = 2^{-\gamma m/400} m!. \tag{2.6.15}$$

Next, we consider the case $i \in [\gamma m/4]$, where first we prove the following claim.

Claim 2.6.3. *For all $i \in [4k]$ we have $|\mathcal{X}^i_{x,y}| \leq (2n)^{i-1}(k+1)i!(m-i)!$.*

Proof. Observe that $|\mathcal{S} \cap \bigcup_{i=1}^{\gamma n/2-1} L_i| \leq 1$. If $\mathcal{S} \cap \bigcup_{i=1}^{\gamma n/2-1} L_i = \varnothing$, then $\mathcal{X}^i_{x,y} = \varnothing$ for all $i \in [4k]$, so assume $|\mathcal{S} \cap \bigcup_{i=1}^{\gamma n/2-1} L_i| = 1$. Let v be the unique vertex in $S \cap \bigcup_{i=1}^{\gamma n/2-1} L_i$. Then,

$B^k_{\mathcal{Q}^n}(v) \cap L_i = B^k_{\mathcal{Q}^n}(\mathcal{S}) \cap L_i$ for each $i \in [4k]$. Thus, in order to prove Claim 2.6.3, it suffices to show that $|B^k_{\mathcal{Q}^n}(v) \cap L_i| \leq (2n)^{i-1}(k+1)$ for each $i \in [4k]$.

We will proceed by induction on i. Since $\varnothing = x \notin B^k_{\mathcal{Q}^n}(v)$, it follows that $|B^k_{\mathcal{Q}^n}(v) \cap L_1| \leq k+1$, so the base case holds.

Now, suppose that $|B^k_{\mathcal{Q}^n}(v) \cap L_{i-1}| \leq (2n)^{i-2}(k+1)$ for some $2 \leq i \leq 4k$. Consider first the case where $v \in L_j$ for some $i \leq j \leq i+k$. In this case, any $u \in L_i \cap B^k_{\mathcal{Q}^n}(v)$ satisfies either

(i) $u \subseteq v$ or

(ii) there is a v-u path of length at most k whose penultimate vertex lies in L_{i-1}.

There are $\binom{j}{i} \leq \binom{k+i}{i}$ choices for u satisfying (i), whereas by applying induction to the penultimate vertex in such paths it follows that there are at most $n(2n)^{i-2}(k+1)$ choices for u satisfying (ii). Altogether, we have

$$|B^k_{\mathcal{Q}^n}(v) \cap L_i| \leq \binom{k+i}{i} + n(2n)^{i-2}(k+1) \leq (2n)^{i-1}(k+1).$$

The case where $v \in L_j$ for some $i-k \leq j < i$ is handled similarly. This completes the induction step and the proof of the claim. ◀

Recall that $|\mathcal{S} \cap B^{\gamma m/2-1}_{\mathcal{Q}^n}(x)| \leq 1$. It follows that for all $i \in [\gamma m/3]$ we have that $|B^k_{\mathcal{Q}^n}(\mathcal{S}) \cap L_i| \leq n^k$ and, therefore, $|\mathcal{X}^i_{x,y}| \leq n^k i!(m-i)!$. Suppose $|\mathcal{S} \cap B^{\gamma m/2-1}_{\mathcal{Q}^n}(x)| = 1$, and let v be the unique vertex in $\mathcal{S} \cap B^{\gamma m/2-1}_{\mathcal{Q}^n}(x)$. Let $j \in [\gamma m/2-1]$ be such that $v \in L_j$. It follows by Claim 2.6.3 that

$$\begin{aligned}\sum_{i=1}^{\gamma m/3} |\mathcal{X}^i_{x,y}| \leq \sum_{i=j-k}^{j+k} |\mathcal{X}^i_{x,y}| &\leq \begin{cases} \sum_{i=j-k}^{j+k} (2n)^{i-1}(k+1)i!(m-i)! & \text{if } j \leq 3k, \\ \sum_{i=j-k}^{j+k} n^k i!(m-i)! & \text{if } 3k < j < \gamma m/2 \end{cases} \\ &\leq \sum_{i=2k}^{4k} (2n)^{i-1}(k+1)i!(m-i)!. \end{aligned} \tag{2.6.16}$$

If $\mathcal{S} \cap B^{\gamma m/2}_{\mathcal{Q}^n}(x) = \varnothing$, then this trivially holds too. By the symmetry of the hypercube, we also have that

$$\sum_{i=m-\gamma m/3}^{m} |\mathcal{X}^i_{x,y}| \leq \sum_{i=2k}^{4k} (2n)^{i-1}(k+1)i!(m-i)!. \tag{2.6.17}$$

Therefore, by (2.6.14)–(2.6.17) we have

$$|\mathcal{X}_{x,y} \setminus \mathcal{X}^{\neg B^k_{\mathcal{Q}^n}(\mathcal{S})}_{x,y}| \leq \sum_{i=\gamma m/3}^{m-\gamma m/3} 2^{-\gamma m/400} m! + 2\sum_{i=2k}^{4k} (2n)^{i-1}(k+1)i!(m-i)! \leq m!/4. \qquad \square$$

Lemma 2.6.16. *Let $0 < 1/n \ll \eta, 1/k', \gamma \leq 1$ and $n/2 \leq k < n$ with $n, k', k \in \mathbb{N}$. Let $\mathcal{S} \subseteq V(\mathcal{Q}^n)$ be such that for all $x \in V(\mathcal{Q}^n)$, we have that $|B^{\gamma n}_{\mathcal{Q}^n}(x) \cap \mathcal{S}| \leq 1$. Let $y \in L_k$ and let $s := \lfloor (k+1)/2 \rfloor$. Then, there exists three sets of vertices $A = \{a_1, \dots, a_{(1-\eta)n}\} \subseteq L_1$, $B = \{b_1, \dots, b_{(1-\eta)n}\} \subseteq N_{Q^n}(y)$ and $C = \{c_1, \dots, c_{(1-\eta)n}\} \subseteq L_s$ such that*

(i) *for each pair $i, j \in [(1-\eta)n]$ with $i \neq j$ we have $\mathrm{dist}(c_i, c_j) \geq 9s^2/(10n)$,*

(ii) *$B^{k'}_{\mathcal{Q}^n}(\mathcal{S}) \cap C = \varnothing$, and*

(iii) *for each $i \in [(1-\eta)n]$ we have $a_i \subseteq c_i \subseteq b_i$.*

Proof. Choose k vertices $c_1, \dots, c_k \in y^{(s)}$ independently and uniformly at random. Then, choose $n-k$ vertices $c'_{k+1}, \dots, c'_n \in y^{(s-1)}$ independently and uniformly at random. For each $i \in [n] \setminus [k]$, choose an element $a_i \in [n] \setminus y$ such that all the a_i are distinct, and let $c_i := c'_i \cup \{a_i\} \in L_s$. For each $i \in [n] \setminus [k]$, let $b_i \in N^{\uparrow}(y)$ be the unique vertex such that $a_i \in b_i$, so that when viewing each a_i now as a 1-element set, we have $a_i \subseteq c_i \subseteq b_i$ for all $i \in [n] \setminus [k]$.

Note that, for each pair $i, j \in [n]$ with $i \neq j$, we have that

$$\mathbb{E}[|c_i \cap c_j|] \leq s^2/k. \tag{2.6.18}$$

Assume that we reveal each c_i in turn. We then have that, for each $i \in [n] \setminus \{1\}$, the variables $|c_i \cap c_j|$ with $j \in [i-1]$ are hypergeometric. Thus, by Lemma 2.4.3 and (2.6.18), for each pair $i, j \in [n]$ with $i \neq j$ we have that

$$\mathbb{P}[|c_i \cap c_j| \geq 21s^2/(20k)] \leq e^{-n/25000}.$$

By a union bound, it follows that a.a.s. for all pairs $i, j \in [n]$ with $i \neq j$ we have $|c_i \cap c_j| < 21s^2/(20k)$ and, thus, $\mathrm{dist}(c_i, c_j) \geq 9s/10$. In particular, (i) holds a.a.s.

Next, let $S_1 := y^{(s)} \cap B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$ and $S_2 := y^{(s-1)} \cap B^{k'+1}_{\mathcal{Q}^n}(\mathcal{S})$. By applying Proposition 2.6.14 first with k', $\mathcal{S}$ and s playing the roles of k, S and t, respectively, and then with $k'+1$, $\mathcal{S}$ and $s-1$ playing the roles of k, S and t, respectively, we obtain that $|S_1| \leq \binom{k}{s}2^{-\gamma n/200}$ and $|S_2| \leq \binom{k}{s-1}2^{-\gamma n/200}$. Therefore, for all $i \in [k]$ we have

$$\mathbb{P}[c_i \in S_1] \leq \frac{\binom{k}{s}2^{-\gamma n/200}}{\binom{k}{s}} = 2^{-\gamma n/200}.$$

For all $i \in [n] \setminus [k]$ we have $\mathbb{P}[c_i \in S_1] \leq \mathbb{P}[c_i' \in S_2]$ and similarly we have $\mathbb{P}[c_i' \in S_2] \leq 2^{-\gamma n/200}$. It now follows by a union bound that (ii) holds a.a.s.

Next, consider an auxiliary bipartite graph H with parts $y^{(1)}$ and $\{c_1, \ldots, c_k\}$ and the following edge set. For each $i \in [k]$ and $a \in y^{(1)}$, let $\{a, c_i\}$ be an edge whenever $a \in c_i$. Thus, for each $i \in [k]$ we have that $d_H(c_i) = s$. Furthermore, it follows by Lemma 2.4.2 that a.a.s. for all $a \in y^{(1)}$ we have $d_H(a) = (1 \pm \eta/2)s$. Condition on this. Then, for each $X \subseteq y^{(1)}$, since we have $e_H(N_H(X), y^{(1)}) \geq e_H(N_H(X), X)$, it follows that $|N(X)| \geq |X| - \eta k/2$. Therefore, by Lemma 2.5.1 we have a matching of size $(1 - \eta/2)k$ in H.

Similarly, a.a.s. we have a matching of size $(1-\eta/2)k$ in the analogous bipartite graph H' with parts $N^{\downarrow}(y)$ and $\{c_1, \ldots, c_k\}$, where for each $i \in [k]$ and $b \in N^{\downarrow}(y)$ we have that $\{b, c_i\}$ is an edge whenever $c_i \subseteq b$. By concatenating these matchings (and relabelling the indices if necessary), it follows that a.a.s. there is an ordering $\{a_1, \ldots, a_k\}$ of the elements of y and an ordering $\{b_1, \ldots, b_k\}$ of the vertices of $N^{\downarrow}(y)$ such that, for all $i \in [(1 - \eta)k]$, we have $a_i \subseteq c_i \subseteq b_i$. Furthermore, as explained before, by construction, for all $i \in [n] \setminus [k]$, we have $a_i \subseteq c_i \subseteq b_i$. Thus, (iii) holds a.a.s.

Finally, given that each of (i), (ii), (iii) holds a.a.s., there must exist a choice of $c_1, \ldots, c_n$ such that (i)-(iii) hold simultaneously. □

We are now in a position to combine the results we have shown so far to prove the following key lemma, which is used to provide a base structure for the near-spanning tree which we seek.

Lemma 2.6.17. *Let* $0 < 1/n \ll 1/C \ll \varepsilon' \leq 1/2$, *and* $0 < 1/n \ll 1/k', \gamma \leq 1/2$, *where* $n, k', C \in \mathbb{N}$, *and let* $(n, \mathbf{p}, M)$ *be feasible with* $0 < 1/n \ll 1/M$. *Moreover, let* $\mathcal{S} \subseteq V(\mathcal{Q}^n)$ *be such that, for all* $x \in V(\mathcal{Q}^n)$, *we have* $|B^{\gamma n}_{\mathcal{Q}^n}(x) \cap \mathcal{S}| \leq 1$ *and* $\varnothing \notin B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$. *Then, with probability at least* $1 - e^{-50n}$ *we have that* $P \sim \mathcal{P}^C(n, \mathbf{p}, M)$ *satisfies the following: for all* $y \in \bigcup_{i=\lceil n/2\rceil}^{\lfloor 9n/10\rfloor} L_i \setminus B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$, *there exists a collection of chains* $\mathcal{X}_y$ *such that, for all* $X \in \mathcal{X}_y$, *we have* $X \subseteq P - B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$, *one of the endpoints of* X *belongs to* L_1, *and*

$$\left|N_{\mathcal{Q}^n}(y) \cap \bigcup_{X \in \mathcal{X}_y} V(X)\right| \geq (1 - \varepsilon')n.$$

Proof. Fix $\eta > 0$ such that $0 < 1/n \ll \eta \ll \varepsilon'$, and let $m := 480000$. Fix a vertex $y \in L_k \setminus B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$ for some $n/2 \leq k \leq 9n/10$. Let $s := \lfloor (k+1)/2 \rfloor$. By Lemma 2.6.16 with $\eta/2$ playing the role of η, there exists a collection of vertices $\{c_1, \ldots, c_{(1-\eta/2)n}\} \subseteq L_s$ such that $B^{k'}_{\mathcal{Q}^n}(c_i) \cap \mathcal{S} = \varnothing$ and $\operatorname{dist}(c_i, c_j) \geq 9s^2/(10n)$ for all pairs $i, j \in [(1-\eta/2)n]$ with $i \neq j$; an ordering $b_1, \ldots, b_n$ of $N_{\mathcal{Q}^n}(y)$, and an ordering $a_1, \ldots, a_n$ of L_1, such that for all $i \in [(1-\eta/2)n]$ we have $a_i \subseteq c_i \subseteq b_i$. For each $i \in [(1-\eta/2)n]$, we call (a_i, b_i, c_i) a *triple*. Note that $|B^{k'}_{\mathcal{Q}^n}(\mathcal{S}) \cap (L_1 \cup N_{\mathcal{Q}^n}(y))| \leq 2(k'+1)$, and hence we may assume for each $i \in [(1-\eta)n]$ that (a_i, b_i, c_i) forms a triple where $a_i, b_i \notin B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$. We denote by $\mathcal{T}$ the collection of all such triples. Partition $[(1-\eta)n]$ into two sets $\mathcal{I}_1 := \{i \in [(1-\eta)n] : b_i \in N^{\downarrow}(y)\}$ and $\mathcal{I}_2 := [(1-\eta)n] \setminus \mathcal{I}_1$. Let $\mathcal{A}_1 := \{a_i : i \in \mathcal{I}_1\}$, $\mathcal{A}_2 := \{a_i : i \in \mathcal{I}_2\}$, $\mathcal{B}_1 := \{b_i : i \in \mathcal{I}_1\}$, $\mathcal{B}_2 := \{b_i : i \in \mathcal{I}_2\}$, $\mathcal{C}_1 := \{c_i : i \in \mathcal{I}_1\}$ and $\mathcal{C}_2 := \{c_i : i \in \mathcal{I}_2\}$. Note that $k - \eta n \leq |\mathcal{C}_1| \leq k$.

We first turn our attention to $\mathcal{A}_1$, $\mathcal{B}_1$ and $\mathcal{C}_1$. Partition $\mathcal{A}_1$, $\mathcal{B}_1$ and $\mathcal{C}_1$ into sets $\mathcal{A}^1, \ldots, \mathcal{A}^m$, $\mathcal{B}^1, \ldots, \mathcal{B}^m$ and $\mathcal{C}^1, \ldots, \mathcal{C}^m$, respectively, each of size at least $\lfloor (k-\eta n)/m \rfloor$ and at most $2\lfloor (k-\eta n)/m \rfloor$, and such that, for every triple $(a, b, c) \in \mathcal{T}$ there exists $j \in [m]$ such that $a \in \mathcal{A}^j$, $b \in \mathcal{B}^j$ and $c \in \mathcal{C}^j$. For each $i \in [m]$, write $\mathcal{A}^i = \{a^i_1, \ldots, a^i_{|\mathcal{A}^i|}\}$, $\mathcal{B}^i = \{b^i_1, \ldots, b^i_{|\mathcal{A}^i|}\}$ and $\mathcal{C}^i = \{c^i_1, \ldots, c^i_{|\mathcal{A}^i|}\}$, where the labeling is such that $(a^i_j, b^i_j, c^i_j) \in \mathcal{T}$ for each $j \in [|\mathcal{A}^i|]$. For each $i \in [m]$ and $j \in [|\mathcal{A}^i|]$, we define the set $\mathcal{Z}_{a^i_j, c^i_j} \subseteq \mathcal{X}_{a^i_j, c^i_j}$ as the set of all chains $X \in \mathcal{X}_{a^i_j, c^i_j}$ which, for all $j' \in [|\mathcal{A}^i|] \setminus \{j\}$, neither intersect any chain $X' \in \mathcal{X}_{a^i_{j'}, c^i_{j'}}$, nor $B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$. By Lemmas 2.6.12 and 2.6.15 and the definition of m, we have that

$$|\mathcal{Z}_{a^i_j,c^i_j}| \geq \frac{1}{2}|\mathcal{X}_{a^i_j,c^i_j}|. \tag{2.6.19}$$

For each triple $(a,b,c) \in \mathcal{T}$ and any graph $G \subseteq \mathcal{Q}^n$, let $I_{a,c}(G)$ take value 1 if $Y(\mathcal{Z}_{a,c},G) > 0$, and 0 otherwise. (Recall that $Y(\mathcal{Z}_{a,c},G)$ denotes the number of chains $X \in \mathcal{Z}_{a,c}$ with $X \subseteq G$.) For each $i \in [m]$, let $I_i(G) \coloneqq \sum_{j\in[|\mathcal{A}^i|]} I_{a^i_j,c^i_j}(G) = \sum_{j\in[|\mathcal{A}^i|]} I_{a^i_j,c^i_j}(G - B^{k'}_{\mathcal{Q}^n}(\mathcal{S}))$.

We are now in a position to consider $P \sim \mathcal{P}^C(n,\mathbf{p},M)$. Recall that P is generated by sampling C independent graphs P_i, where $P_i \sim \mathcal{P}(n,\mathbf{p},M)$. In each P_i we can give bounds on the probability that certain chains appear. Note that, for each $i \in [C]$ and each fixed $i' \in [m]$ we have that, for every pair $j,j' \in [|\mathcal{A}^{i'}|]$ with $j \neq j'$, the variables $Y(\mathcal{Z}_{a^{i'}_j,c^{i'}_j},P_i)$ and $Y(\mathcal{Z}_{a^{i'}_{j'},c^{i'}_{j'}},P_i)$ are independent (and, therefore, $I_{a^{i'}_j,c^{i'}_j}(P_i)$ and $I_{a^{i'}_{j'},c^{i'}_{j'}}(P_i)$ are independent too). Since C is a large constant, this independence will allow us to boost the probability that these chains appear in $P - B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$. The analysis is broken into two steps.

Claim 2.6.4. *With probability at least $1-2e^{-75n}$, the graph $P \sim P^C(n,\mathbf{p},M)$ satisfies the following.*

(1) *$P - B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$ contains an a-c chain for at least $(1-\varepsilon'/2)k$ of the triples $(a,b,c) \in \mathcal{T}$ with $c \in \mathcal{C}_1$.*

(2) *$P - B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$ contains a c-b chain for at least $(1-\varepsilon'/2)k$ of the triples $(a,b,c) \in \mathcal{T}$ with $c \in \mathcal{C}_1$.*

Proof. We show that (1) and (2) each hold with probability $1-e^{-75n}$. The result then follows by a union bound.

For (1), let $C' \coloneqq \sqrt{C}$. By (2.6.19), we can apply Lemma 2.6.11 with $(\varepsilon')^2$, $1/2$ and C' playing the roles of η, α and C, respectively. Thus, for $P' \sim \mathcal{P}^{C'}(n,\mathbf{p},M)$, for all $i \in [m]$ and $j \in [|\mathcal{A}^i|]$ we have that

$$\mathbb{P}[I_{a^i_j,c^i_j}(P') = 1] = \mathbb{P}[Y(\mathcal{Z}_{a^i_j,c^i_j},P') > 0] \geq 1-(\varepsilon')^2.$$

It follows that for all $i \in [m]$ we have $\mathbb{E}[I_i(P')] \geq (1-(\varepsilon')^2)|\mathcal{A}^i|$ and, therefore, by Lemma 2.4.2,

$$\mathbb{P}[I_i(P') > |\mathcal{A}^i|(1-(\varepsilon')^{3/2})] > 1 - e^{-(\varepsilon')^3 n/(25\cdot 10^6)}. \tag{2.6.20}$$

Let $P \sim \mathcal{P}^C(n, \mathbf{p}, M)$, and note that P can be generated by sampling C' independent graphs $P'_j \sim \mathcal{P}^{C'}(n, \mathbf{p}, M)$ and considering their union. For each $i \in [m]$, let $\mathcal{E}^i$ be the event that $I_i(P) > |\mathcal{A}^i|(1-(\varepsilon')^{3/2})$. It follows from (2.6.20) that, for each $i \in [m]$, we have $\mathbb{P}[\mathcal{E}^i] > 1 - e^{-100n}$. Now let $\mathcal{E}$ be the event that, for all $i \in [m]$, $\mathcal{E}_i$ holds. It follows by a union bound that

$$\mathbb{P}[\mathcal{E}] \geq 1 - e^{-75n}.$$

Thus, with probability at least $1 - e^{-75n}$ the graph $P - B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$ contains an a-c chain for at least $(1-(\varepsilon')^{3/2})|\mathcal{C}_1|$ of the triples $(a, b, c) \in \mathcal{T}$ with $c \in \mathcal{C}_1$. Since $|\mathcal{C}_1| \geq (1-2\eta)k$, $P - B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$ contains an a-c chain for at least $(1-\varepsilon'/2)k$ of the triples $(a, b, c) \in \mathcal{T}$ with $c \in \mathcal{C}_1$..

To show (2), for each triple $(a, b, c) \in \mathcal{T}$ with $c \in \mathcal{C}_1$, one can consider the set $\mathcal{X}_{c,b}$ and define sets $\mathcal{Z}_{c,b}$ and variables $I_{c,b}(G)$ analogously to the proof of (1). Then, by Lemma 2.6.11, Lemma 2.6.12 together with Remark 2.6.13, and Lemma 2.6.15, the same argument as above shows that, with probability at least $1 - e^{-75n}$, the graph $P - B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$ contains a c-b chain for at least $(1-\varepsilon'/2)k$ of the triples $(a, b, c) \in \mathcal{T}$ with $c \in \mathcal{C}_1$. ◀

It follows by Claim 2.6.4 that with probability at least $1 - 2e^{-75n}$ we have that $P - B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$ contains an a-b chain for at least $(1-\varepsilon')k$ of the triples $(a, b, c) \in \mathcal{T}$ with $c \in \mathcal{C}_1$. We can prove an analogous result for the sets $\mathcal{A}_2$, $\mathcal{B}_2$ and $\mathcal{C}_2$. More specifically, we can show that with probability at least $1-2e^{-75n}$, for $P \sim P^C(n, \mathbf{p}, M)$, the graph $P - B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$ contains an a-b chain for at least $(1-\varepsilon')(n-k)$ of the triples $(a, b, c) \in \mathcal{T}$ with $c \in \mathcal{C}_2$. Combining this with the previous, it follows that, with probability at least $1 - 4e^{-75n}$, $P - B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$ contains an a-b chain for at least $(1-\varepsilon')n$ of the triples $(a, b, c) \in \mathcal{T}$. Finally, the result follows by a union bound over all $y \in \bigcup_{i=\lceil n/2 \rceil}^{\lfloor 9n/10 \rfloor} L_i \setminus B^{k'}_{\mathcal{Q}^n}(\mathcal{S})$. □

Let F be the union of all chains given by Lemma 2.6.17 (applied with $k' \coloneqq k$). Then, F satisfies (T2) in Theorem 2.6.1 for all vertices $x \in \bigcup_{i=\lceil n/2\rceil}^{\lfloor 9n/10\rfloor} L_i \setminus B^k_{\mathcal{Q}^n}(\mathcal{S})$. However, we need this property to hold for every $x \in V(\mathcal{Q}^n) \setminus B^k_{\mathcal{Q}^n}(\mathcal{S})$. Recall the discussion in the beginning of this section where, due to the symmetries in the hypercube, we can 'redefine' any vertex $v \in V(\mathcal{Q}^n)$ to be the empty set $\varnothing$. As discussed, this leads to a redefined notion of levels in the hypercube where, for each $i \in [n]_0$, we let $L_i(v) \coloneqq \{u \in V(\mathcal{Q}^n) : \mathrm{dist}(u,v) = i\}$. The notion of a chain in this setting was also discussed.

When we consider this generalised setting, by replacing L_i with $L_i(v)$ in Definitions 2.6.2, 2.6.3 and 2.6.10, we obtain a distribution on subgraphs of $\mathcal{Q}^n$ which we denote by $\mathcal{P}^C_v(n,\mathbf{p},M)$. (Note, again, that there is a joint distribution of $\mathcal{P}^C_v(n,\mathbf{p},M)$ and $\mathcal{Q}^n_{\min\{1,Cp\}}$ such that $\mathcal{P}^C_v(n,\mathbf{p},M) \subseteq Q^n_{\min\{1,Cp\}}$, where $p = \max_{i\in[n-1]_0} p_i$.) Then, for any fixed $v \in V(\mathcal{Q}^n)$, Lemma 2.6.17 holds in this setting by replacing chains by chains with respect to v. Intuitively, we may think of this simply as growing several branching processes rooted at different vertices of the hypercube. This will be crucial in proving (T2).

Note that F may have unbounded degrees and also may be disconnected. To turn F into a bounded degree forest we will later delete suitable edges. To make it connected without significantly raising any vertex degrees we will apply the following lemma.

Lemma 2.6.18. *For $n \in \mathbb{N}$ such that $0 < 1/n \ll \delta \leq 1/50$ and $0 < \varepsilon < 1/2$, the following holds a.a.s. Let $R \sim Res(\mathcal{Q}^n, \delta)$. Then, there exists a cycle in $\mathcal{Q}^n_\varepsilon[(L_1 \cup L_2) \setminus R]$ which covers $L_1 \setminus R$.*

Proof. Let $R \sim Res(\mathcal{Q}^n, \delta)$. Let $\mathcal{A}$ be the event that $|R \cap L_1| \geq n/4$. By Lemma 2.4.4 we have that $\mathbb{P}[\mathcal{A}] \leq e^{-\Theta(n)}$. Expose $R \cap L_1$ and condition on the event that $\mathcal{A}$ does not occur.

Note that for each pair of vertices $x, y \in L_1$ there exists a unique vertex $z \in L_2 \cap N_{\mathcal{Q}^n}(x) \cap N_{\mathcal{Q}^n}(y)$ (in particular, $z = x \cup y$). Let H be an auxiliary graph with vertex set $L_1 \setminus R$, where we include an edge between x and y if $x \cup y \notin R$ and $\{x, x\cup y\}, \{y, x \cup y\} \in E(\mathcal{Q}^n_\varepsilon)$. By definition, a Hamilton cycle in H would correspond uniquely to a cycle in $\mathcal{Q}^n_\varepsilon[(L_1 \cup L_2) \setminus R]$

covering $L_1 \setminus R$. Note that H has the same distribution as a binomial random graph $G \sim G_{n-|R\cap L_1|,p}$, where $p=(1-\delta)\varepsilon^2$. Let $\mathcal{B}$ be the event that there exists a Hamilton cycle in H. As, after conditioning on $\mathcal{A}$ not holding, $G_{n-|R\cap L_1|,p}$ is a.a.s. Hamiltonian (see e.g. [73, 87]), it follows that

$$\mathbb{P}[\mathcal{B}] \geq \mathbb{P}[\mathcal{B} \mid \overline{\mathcal{A}}]\mathbb{P}[\overline{\mathcal{A}}] \geq (1-o(1))(1-e^{-\Theta(n)}) = 1-o(1). \qquad \square$$

We are now in a position to prove Theorem 2.6.1.

Proof of Theorem 2.6.1. Choose constants $M, C \in \mathbb{N}$ such that $1/D, \delta \ll 1/C, 1/M \ll \varepsilon'$. By Proposition 2.6.7, there exists a tuple $(n, \mathbf{p}, M)$ which is feasible and such that $\max_{i\in[n-1]_0} p_i \leq \varepsilon/(5C)$. Let $x_1 \coloneqq \varnothing$, $x_2 \coloneqq [\lceil n/2\rceil]$, $x_3 \coloneqq [n]\setminus x_2$ and $x_4 \coloneqq [n]$. For each $j\in[4]$, let $P_j \sim \mathcal{P}^C_{x_j}(n,\mathbf{p},M)$ be sampled independently, and let R_j be the reservoir associated with P_j. Let $R \coloneqq \bigcap_{j\in[4]} R_j$, and note that $R \sim \mathit{Res}(\mathcal{Q}^n, 1/10^{8C})$. Finally, let $Q \sim \mathcal{Q}^n_{\varepsilon/5}$ be independent of all other previous choices. Recall that, for each $j\in[4]$, there is a joint distribution of $\mathcal{P}^C_{x_j}(n,\mathbf{p},M)$ and $\mathcal{Q}^n_{\varepsilon/5}$ such that $\mathcal{P}^C_{x_j}(n,\mathbf{p},M) \subseteq \mathcal{Q}^n_{\varepsilon/5}$ (see the discussion after Definition 2.6.10). It follows that there is a joint distribution of $\bigtimes_{j=1}^4 \mathcal{P}^C_{x_j}(n,\mathbf{p},M) \times \mathcal{Q}^n_{\varepsilon/5}$ and $\mathcal{Q}^n_\varepsilon$ such that $P_1\cup P_2\cup P_3\cup P_4\cup Q \subseteq \mathcal{Q}^n_\varepsilon$. Therefore, it suffices to show that we can find the desired tree T in $(P_1\cup P_2\cup P_3\cup P_4\cup(Q-R)) - B^k_{\mathcal{Q}^n}(\mathcal{S})$.

For each $j\in[4]$, let $A_j \coloneqq \bigcup_{i=\lceil n/2\rceil}^{\lfloor 9n/10\rfloor} L_i(x_j)\setminus B^k_{\mathcal{Q}^n}(\mathcal{S})$, and let $\mathcal{E}_j$ be the event that, for all $y\in A_j$, the graph $P_j - B^k_{\mathcal{Q}^n}(\mathcal{S})$ contains a collection $\mathcal{X}^j_y$ of chains with respect to x_j, where each chain $X\in\mathcal{X}^j_y$ has an endpoint in $L_1(x_j)$ (and thus in $L_1(x_j)\setminus(R_j\cup B^k_{\mathcal{Q}^n}(\mathcal{S}))$), and such that at least $(1-\varepsilon')n$ of the neighbours of y in $\mathcal{Q}^n$ are covered by the union of the chains in $\mathcal{X}^j_y$. Note that $\mathcal{E}_j$ is equivalent to saying that the union of the chains in $\mathcal{X}^j_y$ satisfies (T2) for all $y\in A_j$. For each $j\in[4]$ we have by Lemma 2.6.17 that $\mathbb{P}[\mathcal{E}_j]\geq 1-e^{-50n}$. Condition on the event that $\mathcal{E}_j$ holds for all $j\in[4]$.

For each $j\in[4]$, let $F_j\subseteq\mathcal{Q}^n$ be given by $F_j \coloneqq \bigcup_{y\in A_j}\bigcup_{X\in\mathcal{X}^j_y} X$. For each $j\in[4]$, let $G_j\subseteq F_j$ be defined by removing, for each $y\in V(F_j)\setminus\{x_j\}$, all edges of F_j joining y to its down-neighbours with respect to x_j except for one (if y has one such down-neighbour in

F_j). In particular, it follows that each connected component of G_j is a tree and contains one vertex in $L_1(x_j)$, and that $\Delta(G_j) \leq CM+1$. Since G_j has the same vertex set as F_j, we have that G_j satisfies (T2) for all $y \in A_j$. Furthermore, note that $V(\mathcal{Q}^n) \setminus B^k_{\mathcal{Q}^n}(\mathcal{A}) = \bigcup_{j=1}^4 A_j$. Therefore, the graph $G \coloneqq \bigcup_{j\in[4]} G_j$ satisfies (T2) and $\Delta(G) \leq 4CM+4$.

Since $B^{k+2}_{\mathcal{Q}^n}(\mathcal{S}) \cap \{\varnothing, [n], [\lceil n/2 \rceil], [n] \setminus [\lceil n/2 \rceil]\} = \varnothing$ it follows that $B^k_{\mathcal{Q}^n}(\mathcal{S}) \cap (L_1(x_j) \cup L_2(x_j)) = \varnothing$ for each $j \in [4]$. Let $\mathcal{E}_5$ be the event that, for each $j \in [4]$, $Q[L_1(x_j) \cup L_2(x_j)] - R$ contains a cycle C_j which covers $L_1(x_j) \setminus R$. By four applications of Lemma 2.6.18 (applied with x_j playing the role of $\varnothing$) we have that $\mathbb{P}[\mathcal{E}_5] = 1 - o(1)$. Condition on the event that this occurs.

Let $H \coloneqq G \cup \bigcup_{j\in[4]} C_j$. It follows that H is connected and $\Delta(H) \leq 4CM+6$. In order to complete the proof, let $T \subseteq H$ be a spanning tree of H. □

2.6.2 Extending the tree

Roughly speaking, in Theorem 2.6.1 we showed that, for any $\varepsilon > 0$, given a reservoir chosen at random, the random graph $\mathcal{Q}^n_\varepsilon$ a.a.s. contains a bounded-degree tree T' which avoids the reservoir and satisfies the local property that, for every vertex $x \in V(\mathcal{Q}^n)$, all but a fixed small proportion of its neighbours are covered by T'. Our goal in this section is to show that T' can be extended into a tree T where the proportion of uncovered vertices (in each neighbourhood) is even smaller, while still retaining the bounded degree property. The precise statement is the following.

Theorem 2.6.19. *For all $0 < 1/n \ll 1/\ell, \varepsilon \leq 1$, where $n, \ell \in \mathbb{N}$, the following holds. Let $R, W \subseteq V(\mathcal{Q}^n)$ and let $T' \subseteq \mathcal{Q}^n - (R \cup W)$ be a tree. For each $x \in V(\mathcal{Q}^n) \setminus W$, let $Z(x) \subseteq N_{\mathcal{Q}^n}(x) \cap V(T')$ be such that $|Z(x)| \geq 3n/4$. Then, a.a.s. there exists a tree T with $T' \subseteq T \subseteq (\mathcal{Q}^n_\varepsilon \cup T') - W$ such that*

(TC1) $\Delta(T) \leq \Delta(T') + 1$;

(TC2) *for all $x \in V(\mathcal{Q}^n)$, we have that $|B^\ell_{\mathcal{Q}^n}(x) \setminus (V(T) \cup W)| \leq n^{3/4}$, and*

(TC3) *for each* $x \in V(T) \cap R$, *we have that* $d_T(x) = 1$ *and the unique neighbour* x' *of* x *in* T *is such that* $x' \in Z(x)$.

Proof. Let $Q \sim \mathcal{Q}^n_\varepsilon$. For each $x \in V(\mathcal{Q}^n) \setminus W$ we have $3\varepsilon n/4 \leq \mathbb{E}[e_Q(x, Z(x))] \leq \varepsilon n$. Let $S_1 := \{x \in V(\mathcal{Q}^n) : d_Q(x) > 11\varepsilon n/10\}$, $S_2 := \{x \in V(\mathcal{Q}^n) \setminus W : e_Q(x, Z(x)) < 2\varepsilon n/3\}$ and $S := S_1 \cup S_2$. Let $\mathcal{E}_1$ be the event that there exists no vertex $x \in V(\mathcal{Q}^n)$ such that $|B^\ell_{\mathcal{Q}^n}(x) \cap S_1| \geq n^{1/2}$. By Lemma 2.5.6 we have that $\mathbb{P}[\mathcal{E}_1] \geq 1 - e^{-4n}$. Similarly, let $\mathcal{E}_2$ be the event that there exists no vertex $x \in V(\mathcal{Q}^n)$ such that $|B^\ell_{\mathcal{Q}^n}(x) \cap S_2| \geq n^{1/2}$. By Lemma 2.5.6 we have that $\mathbb{P}[\mathcal{E}_2] \geq 1 - e^{-4n}$. Condition on $\mathcal{E}_1 \wedge \mathcal{E}_2$ holding, that is, that there is no vertex $x \in V(\mathcal{Q}^n)$ such that $|B^\ell_Q(x) \cap S| \geq n^{3/4}$.

Given $\mathcal{E}_1 \wedge \mathcal{E}_2$, let H be an auxiliary bipartite graph with parts $A := V(T') \setminus S$ and $B := V(\mathcal{Q}^n) \setminus (V(T') \cup W \cup S)$, where we include an edge between $a \in A$ and $b \in B$ whenever $\{a, b\} \in E(Q)$ and $a \in Z(b)$. By definition of S we have for all $a \in A$ that

$$d_H(a) \leq 11\varepsilon n/10 - 2\varepsilon n/3 < \varepsilon n/2.$$

Furthermore, we have for all $b \in B$ that

$$d_H(b) \geq 2\varepsilon n/3 - n^{3/4} > \varepsilon n/2.$$

Since for all $X \subseteq B$ we have $e_H(N_H(X), B) \geq e_H(X, N_H(X))$, it follows that $|N_H(X)| \geq |X|$. Thus, by Lemma 2.5.1, H contains a matching covering all of B. This corresponds to a matching in $\mathcal{Q} \sim \mathcal{Q}^n_\varepsilon$. The statement follows by setting T to be the union of T' and this matching. □

2.6.3 The repatching lemma

Later we will apply Theorem 2.6.1 to obtain a tree T and a reservoir R in $\mathcal{Q}^n_\varepsilon$ which is disjoint from $V(T)$. To carry out the absorption step later on, it will be important that for each vertex some proportion of its neighbourhood consists of vertices in R. However, the tree produced by Theorem 2.6.1 (and the subsequent application of Theorem 2.6.19) will result in a small number of vertices with few or no neighbours in R. The following repatching lemma will be called on to deal with such vertices, by slightly altering T.

Given a graph P and $S \subseteq V(P)$ we say that S *is connected in* P if the vertices of S lie in the same component of P.

Lemma 2.6.20. *Let $0 < 1/n \ll c, \varepsilon, 1/f, 1/D$ where $f, D \in \mathbb{N}$. Given a fixed $x \in V(\mathcal{Q}^n)$, let $C(x) \subseteq N_{\mathcal{Q}^n}(x) \times N_{\mathcal{Q}^n}(x)$ be such that $|C(x)| \geq cn$ and such that, for all distinct $(y_1, z_1), (y_2, z_2) \in C(x)$, we have $\{y_1, z_1\} \cap \{y_2, z_2\} = \varnothing$. Furthermore, for each $(y, z) \in C(x)$, let $B(y, z) \subseteq (N_{\mathcal{Q}^n}(y) \cup N_{\mathcal{Q}^n}(z)) \setminus \{x\}$ with $|B(y, z)| < D$. Then, with probability at least $1 - e^{-5n}$, for every $F \subseteq V(\mathcal{Q}^n)$ with $|F| \leq f$, there exist a pair $(y, z) \in C(x)$ with $y, z \notin F$ and a graph $P \subseteq \mathcal{Q}^n_\varepsilon - \{y, z\}$ with $|V(P)| < 5D$ such that*

(R1) *$B(y, z) \cap N_{\mathcal{Q}^n}(y)$ is connected in P, and so is $B(y, z) \cap N_{\mathcal{Q}^n}(z)$.*

(R2) *$V(P) \cap F = \varnothing$.*

Proof. We provide a counting argument to show there exist edge-disjoint graphs $P_1, \ldots, P_{\varepsilon' n} \subseteq \mathcal{Q}^n$ such that, if any is present in $\mathcal{Q}^n_\varepsilon$, then it would satisfy (R1) and (R2) for some $(y, z) \in C(x)$. We will then prove that, with high probability, one of the P_i must be present in $\mathcal{Q}^n_\varepsilon$. Note that we may assume $x = \varnothing$. By passing to a subset of $C(x)$ and replacing c with $c/(30D)$ if necessary, we may also assume that $|C(x)| = cn$ and $2Dc < 1/10$. Similarly, by passing to a suitable subset of $C(x)$, we may assume that, for any distinct $(y, z), (y', z') \in C(x)$, we have that $B(y, z) \cap B(y', z') = \varnothing$.

Fix any $F \subseteq V(\mathcal{Q}^n)$ with $|F| \leq f$. We update $C(x)$ by removing any pair $(y, z) \in C(x)$ for which $(\{y, z\} \cup B(y, z)) \cap F \neq \varnothing$. It follows that $|C(x)| \geq cn - 2f$. Now, for each $(y, z) \in C(x)$ and for each $w \in \{y, z\}$, let $A_w := N_{\mathcal{Q}^n}(w) \cap B(y, z)$, and let $x^w_1, \ldots, x^w_{|A_w|}$ be the vertices of A_w.

Claim 2.6.5. *For each $e = (y, z) \in C(x)$, $w \in \{y, z\}$ and $i \in [|A_w| - 1]$, there exists a collection $\mathcal{P}^w_i$ of subgraphs of $\mathcal{Q}^n$ such that the following hold:*

(RC1) $|\mathcal{P}_i^w| \geq n/2$ *and for each* $P \in \mathcal{P}_i^w$ *we have* $V(P) \cap (F \cup \{y, z\}) = \varnothing$.

(RC2) *Every* $P \in \mathcal{P}_i^w$ *is an* (x_i^w, x_{i+1}^w)*-path of length* 4.

(RC3) *The graphs in* $\mathcal{P}_i^w$ *are pairwise edge-disjoint.*

(RC4) *For every* $e' = (y', z') \in C(x)$ *with* $e' \neq e$, *every* $w' \in \{y', z'\}$ *and every* $j \in [|A_{w'}| - 1]$, *the graphs in* $\mathcal{P}_i^w$ *are edge-disjoint from those in* $\mathcal{P}_j^{w'}$.

(Note that we do not require the paths in $\mathcal{P}_i^w$ to be edge-disjoint from those in $\mathcal{P}_{i'}^{w'}$ when $w, w' \in \{y, z\}$ are distinct and $i \in [|A_w| - 1]$, $i' \in [|A_{w'}| - 1]$.)

Proof of Claim 2.6.5. Let $e_1, \ldots, e_{cn}$ be an ordering of the elements of $C(x)$, where for each $k \in [cn]$ we have that $e_k = (y_k, z_k)$. Note that, for each $k \in [cn]$, each $w \in \{y_k, z_k\}$ and all $i \in [|A_w|]$, we have that $|x_i^w| = 2$, and for each $i, j \in [|A_w|]$ with $i \neq j$ we have that $\mathrm{dist}(x_i^w, x_j^w) = 2$, with $x_i^w \cap x_j^w = w$.

Suppose that, for some $1 < k \leq cn$, every $j \in [k-1]$, every $w \in \{y_j, z_j\}$ and every $i \in [|A_w| - 1]$, we have found a collection $\mathcal{P}_i^w$ which satisfies (RC1)–(RC4). We now show that, for each $w \in \{y_k, z_k\}$ and each $i \in [|A_w| - 1]$, a suitable choice for $\mathcal{P}_i^w$ exists. We construct the set $\mathcal{P}_i^w$ as follows. Let $v_1 := x_i^w \setminus w$ and $v_2 := x_{i+1}^w \setminus w$. For each $d \in [n] \setminus (x_i^w \cup x_{i+1}^w)$, let $P_d \subseteq \mathcal{Q}^n$ be the path which passes through the following vertices in successive order:

$$x_i^w, x_i^w \cup \{d\}, x_i^w \cup \{d\} \cup v_2, (x_i^w \cup \{d\} \cup v_2) \setminus v_1 = x_{i+1}^w \cup \{d\}, x_{i+1}^w.$$

Note that each path P_d has length 4 and that $V(P_d) \cap \{y_k, z_k\} = \varnothing$. Furthermore, for any distinct $d, d' \in [n] \setminus (x_i^w \cup x_{i+1}^w)$, it is clear that P_d and $P_{d'}$ are internally disjoint, and hence, are edge-disjoint. To avoid F as well as the edges of any previously chosen paths we set

$$\mathcal{P}_i^w := \left\{ P_d : d \in [n] \setminus (x_i^w \cup x_{i+1}^w); x_i^w \cup \{d\}, x_{i+1}^w \cup \{d\} \notin N\left(\bigcup_{j=1}^{k-1} B(y_j, z_j) \right); V(P_d) \cap F = \varnothing \right\}.$$

It follows that $\mathcal{P}_i^w$ satisfies (RC2) and (RC3). Recall that $V(P_d) \cap \{y_k, z_k\} = \varnothing$. Therefore, to see that (RC1) holds, note that, for all distinct $(y', z'), (y'', z'') \in C(x)$ and all

$x' \in B(y',z'), x'' \in B(y'',z'')$, since $B(y',z') \cap B(y'',z'') = \varnothing$, we have that x' and x'' contain at most one common neighbour in the third level L_3 of $\mathcal{Q}^n$. Since $|\bigcup_{j=1}^{k-1} B(y_j,z_j)| < Dcn$, there are at most $2Dcn < n/10$ choices for d such that $x_i^w \cup \{d\} \in N(\bigcup_{j=1}^{k-1} B(y_j,z_j))$, or $x_{i+1}^w \cup \{d\} \in N(\bigcup_{j=1}^{k-1} B(y_j,z_j))$. Furthermore, since $|F| \leq f$, it follows that there are still at least $n/2$ suitable choices for d, that is, (RC1) holds as desired. Additionally, (RC4) holds by construction; indeed, since neither the second nor the fourth vertex of each path in $\mathcal{P}_i^w$ lies in some path in $\bigcup_{j\in[k-1]}\bigcup_{w'\in\{y_j,z_j\}}\bigcup_{i'\in[|A_{w'}|-1]}\mathcal{P}_{i'}^{w'}$, the paths in $\mathcal{P}_i^w$ must be edge-disjoint from all the paths in $\bigcup_{j\in[k-1]}\bigcup_{w'\in\{y_j,z_j\}}\bigcup_{i'\in[|A_{w'}|-1]}\mathcal{P}_{i'}^{w'}$. Thus, we can proceed by induction and create a suitable collection $\mathcal{P}_i^w$ for each $k \in [cn]$, $w \in \{y_k, z_k\}$ and $i \in [|A_w|-1]$. ◀

For each $e = (y,z) \in C(x)$, $w \in \{y,z\}$ and $i \in [|A_w|-1]$, let $\mathcal{P}_i^w$ be the collection of subgraphs given by Claim 2.6.5. Note that, for any choice of $P_1 \in \mathcal{P}_1^w, \dots, P_{|A_w|-1} \in \mathcal{P}_{|A_w|-1}^w$, we have that A_w is connected in $P_w \coloneqq \bigcup_{j=1}^{|A_w|-1} P_j$. To complete the proof, we now show that, on passing to $\mathcal{Q}_\varepsilon^n$, with high probability there will exist some $e = (y,z) \in C(x)$ and some P_y and P_z of the above form such that $P_y \cup P_z \subseteq \mathcal{Q}_\varepsilon^n$. Moreover, note that each such choice of $P_y \cup P_z$ satisfies (R1) and (R2) for our fixed F and $|P_y \cup P_z| \leq 5D$. Since $P_y \cup P_z \subseteq B_{\mathcal{Q}^n}^4(x)$, Lemma 2.6.20 will then follow by a union bound over all choices of $F \subseteq B_{\mathcal{Q}^n}^4(x)$ with $|F| \leq f$.

Let $Q \sim \mathcal{Q}_\varepsilon^n$. Consider $e = (y,z) \in C(x), w \in \{y,z\}$ and $i \in [|A_w|-1]$. Let $P \in \mathcal{P}_i^w$ and recall that P has length 4. It follows that $\mathbb{P}[P \nsubseteq Q] = 1 - \varepsilon^4$. Let $\mathcal{E}_i^w$ be the event that there exists some $P \in \mathcal{P}_i^w$ such that $P \subseteq Q$. Since $|\mathcal{P}_i^w| \geq n/2$ and paths in $\mathcal{P}_i^w$ are edge-disjoint by (RC3), we have that $\mathbb{P}[\mathcal{E}_i^w] \geq 1-(1-\varepsilon^4)^{n/2}$. Let $\mathcal{E}_e \coloneqq \bigwedge_{w\in\{y,z\}}\bigwedge_{i\in[|A_w|-1]}\mathcal{E}_i^w$. Since $|A_y| + |A_z| \leq 2D$, we have that

$$\mathbb{P}[\mathcal{E}_e] \geq 1 - 2D(1-\varepsilon^4)^{n/2} \geq 1 - e^{-\varepsilon^4 n/4}.$$

Finally, let $\mathcal{E}$ be the event that there exists some $e \in C(x)$ such that the event $\mathcal{E}_e$ occurs. It follows by (RC4) that, for $e, e' \in C(x)$ with $e \neq e'$, the event $\mathcal{E}_e$ is independent of $\mathcal{E}_{e'}$. Therefore, since $|C(x)| \geq cn$, we have that

$$\mathbb{P}[\mathcal{E}] \geq 1 - e^{-\varepsilon^4 cn^2/4}.$$

Recall that by (RC2) it now suffices to consider a union bound over all choices of $F \subseteq B^4_{\mathcal{Q}^n}(x)$ with $|F| \leq f$. The result follows since

$$1 - f\binom{n^4}{f}e^{-\varepsilon^4 cn^2/4} > 1 - e^{-5n}. \qquad \square$$

2.7 Hamilton cycles in randomly perturbed dense subgraphs of the hypercube

In this section, we introduce a few more auxiliary lemmas and combine them with the tools we have developed so far to prove the following result.

Theorem 2.7.1. *For every $\varepsilon, \alpha \in (0, 1]$ and $c > 0$, there exists $\Phi \in \mathbb{N}$ such that the following holds. Let $H \subseteq \mathcal{Q}^n$ be a spanning subgraph with $\delta(H) \geq \alpha n$ and let $G \sim \mathcal{Q}^n_\varepsilon$. Then, a.a.s. there is a subgraph $G' \subseteq G$ with $\Delta(G') \leq \Phi$ such that, for every $F \subseteq \mathcal{Q}^n$ with $\Delta(F) \leq c\Phi$, the graph $((H \cup G) \setminus F) \cup G'$ is Hamiltonian.*

Note that Theorem 2.7.1 trivially implies the case $k = 1$ of Theorem 2.1.7. In fact, in Section 2.7.5 we will use Theorem 2.7.1 to prove Theorem 2.1.7 in full generality. For this derivation, we will need the stronger conditions imposed in the statement of Theorem 2.7.1. More precisely, the formulation of Theorem 2.7.1 involving a 'forbidden' graph F and a 'protected' graph G' is designed to make repeated applications of Theorem 2.7.1 possible in order to take out k edge-disjoint Hamilton cycles. When finding the i-th Hamilton cycle, the protected graph will contain all the essential ingredients for this, while the forbidden graph will contain all previously chosen Hamilton cycles as well as the protected graphs for the entire set of Hamilton cycles (see Section 2.7.5 for details).

The first step of the proof of Theorem 2.7.1 will be to consider a particular partition of the hypercube into subcubes. The structure of this partition will be used extensively throughout the rest of the paper, so we first introduce the necessary notation in the next subsection. Then, in Section 2.7.2 we prove several results regarding this structure, concerning its properties in $\mathcal{Q}^n_\varepsilon$ and with respect to a reservoir $R \sim Res(\mathcal{Q}^n, \delta)$. In Section 2.7.3, we will prove our *connecting lemmas*, which provide sets of paths in (sub)cubes which (roughly speaking) link up pairs of vertices and, together, span all vertices of these (sub)cubes. We prove Theorem 2.7.1 in Section 2.7.4. Finally, we deduce Theorems 2.1.1, 2.1.2 and 2.1.7 from Theorem 2.7.1 in Section 2.7.5.

2.7.1 Layers, molecules, atoms and absorbing structures.

Throughout this section, given any two vectors u and v, we will write uv for their concatenation. Consider $\mathcal{Q}^n$ and some $s \in \mathbb{N}$, with $s < n$. We divide $\mathcal{Q}^n$ into 2^s vertex-disjoint copies of $\mathcal{Q}^{n-s}$ as follows: for each $a \in \{0,1\}^s$, we consider the set of vertices $V_a := \{av : v \in \{0,1\}^{n-s}\}$, and consider the graph $\mathcal{Q}(a) := \mathcal{Q}^n[V_a]$. We will refer to each $\mathcal{Q}(a)$ as an *s-layer* of $\mathcal{Q}^n$ (s will be dropped whenever clear from the context). Given $\ell \leq n-s$, we will refer to any copy of a cube $\mathcal{Q}^\ell$ in one of the s-layers as an *ℓ-atom* (again, ℓ will be dropped whenever clear from the context).

Fix a Hamilton cycle $\mathcal{C}$ of $\mathcal{Q}^s$. By abusing notation, whenever necessary, we assume that the coordinate vector of each vertex of $\mathcal{C}$ is concatenated with $n-s$ 0's. $\mathcal{C}$ induces a cyclical ordering on $\{0,1\}^s$, which we will label as $a_1, \ldots, a_{2^s}$. In turn, this gives a cyclical ordering on the set of layers. In this section, for each $i \in [2^s]$, we denote $L_i := \mathcal{Q}(a_i)$ (as opposed to Section 2.6, where L_i denoted the i-th level of the hypercube). Given an ℓ-atom $\mathcal{A}$ in an s-layer $\mathcal{Q}(a)$, we refer to $\mathcal{M}(\mathcal{A}) := \mathcal{A} + V(\mathcal{C})$ as an *(s,ℓ)-molecule* (again, the parameters will be dropped when clear from the context). Thus $\mathcal{M}(\mathcal{A})$ is the vertex-disjoint union of 2^s copies of $\mathcal{Q}^\ell$. We refer to an $(s,1)$-molecule as a *vertex molecule* and an (s,ℓ)-molecule for $\ell > 1$ as a *cube molecule*. Observe that, if we label the atoms in a molecule cyclically following the labelling of the layers, then $\mathcal{Q}^n$ contains a perfect matching between any two consecutive atoms where all edges are in the same direction as the corresponding edge in $\mathcal{C}$. Whenever we work with molecules, we consider this cyclical order implicitly. In particular, whenever we refer to a molecule $\mathcal{M} = \mathcal{M}(\mathcal{A}) = \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_{2^s}$, the cyclical order $\mathcal{A}_1 \cup \cdots \cup \mathcal{A}_{2^s}$ of the $\mathcal{A}_i$ is that induced by $\mathcal{C}$. Given a molecule $\mathcal{M}(\mathcal{A}) = \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_{2^s}$, a *slice* $\mathcal{M}^* \subseteq \mathcal{M}(\mathcal{A})$ will consist of the subgraph of $\mathcal{M}(\mathcal{A})$ induced by its intersection with some number of consecutive layers, i.e. $\mathcal{M}^* = \mathcal{A}_{a+1} \cup \cdots \cup \mathcal{A}_{a+t}$ for some $a, t \in [2^s]$. Alternatively, given any $a \in V(\mathcal{C})$, any path $P \subseteq \mathcal{C}$ and any atom $\mathcal{A} \subseteq \mathcal{Q}(a)$, P determines a slice of $\mathcal{M}(\mathcal{A})$ by setting $\mathcal{M}^* := \mathcal{A} + V(P)$.

Consider $i \in [2^s]$ and the cyclical ordering of the layers given by $\mathcal{C}$. Given any subgraph $G \subseteq \mathcal{Q}^n$, we will often denote the restriction of G to the i-th layer by $L_i(G)$, that is, $L_i(G) := G[V(L_i)]$. Given any $v \in \{0,1\}^{n-s}$, we will refer to the vertex $a_i v$ as the i-th *clone* of v. In general, when it is clear from the context, we will also refer to the i-th clone of a cube $C \subseteq \mathcal{Q}^{n-s}$ (as well as other subgraphs), which, analogously, will be the corresponding copy in L_i of C. In particular, the i-th layer L_i is the i-th clone of $\mathcal{Q}^{n-s}$.

As we already discussed in Section 2.2, in order to prove our results we will first construct a near-spanning cycle and then absorb the remaining vertices into this cycle. We will achieve this by using the following absorbing structure.

Definition 2.7.2 (Absorbing ℓ-cube pair)**.** *Let $\ell, n \in \mathbb{N}$, and let $G \subseteq \mathcal{Q}^n$. Given a vertex $x \in V(\mathcal{Q}^n)$, an* absorbing ℓ-cube pair *for x in G, which we denote by (C^l, C^r), is a subgraph of G which consists of two vertex-disjoint ℓ-dimensional cubes $C^l, C^r \subseteq G$ and three edges $e, e^l, e^r \in E(G)$ satisfying the following properties:*

(AP1) $|V(C^l) \cap N_{\mathcal{Q}^n}(x)| = |V(C^r) \cap N_{\mathcal{Q}^n}(x)| = 1$*;*

(AP2) *e^l and e^r are the unique edges from x to C^l and C^r, respectively;*

(AP3) *the unique vertex $y \in V(C^l) \cap N_{\mathcal{Q}^n}(x)$ satisfies* $\mathrm{dist}(y, C^r) = 1$*, and*

(AP4) *e is the unique edge from y to C^r.*

We will refer to C^l as the left absorption cube *and to C^r as the* right absorption cube*. Given an absorbing ℓ-cube pair (C^l, C^r) we refer to y as the* left absorber tip*, and to the* *unique vertex* $z \in V(C^r) \cap N_{\mathcal{Q}^n}(x)$ *as the* right absorber tip. *We refer to the unique vertex* $z' \in e \setminus \{y\}$ *as the* third absorber vertex.

2.7.2 Bondless and bondlessly surrounded molecules

Given any graph $G \subseteq \mathcal{Q}^n$, we will say that an (s,ℓ)-molecule $\mathcal{M} = \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_{2^s} \subseteq \mathcal{Q}^n$, where $\mathcal{A}_i$ is the i-th clone of some ℓ-cube $\mathcal{A} \subseteq \mathcal{Q}^{n-s}$, is *bonded* in G if, for all $i \in [2^s]$, G contains at least 100 edges between $\mathcal{A}_i$ and $\mathcal{A}_{i+1}$ whose endpoint in $\mathcal{A}_i$ has even parity and at least 100 such edges whose endpoint in $\mathcal{A}_i$ has odd parity. Otherwise, we call it *bondless* in G. Furthermore, given a collection $\mathcal{U}$ of (s,ℓ)-molecules in G, we say that $\mathcal{M} \in \mathcal{U}$ is *bondlessly surrounded* in G (with respect to $\mathcal{U}$) if there exists some vertex $v \in V(\mathcal{M})$ which has at least $n/2^{\ell+5s}$ neighbours in $\mathcal{Q}^n$ which are part of (s,ℓ)-molecules of $\mathcal{U}$ which are bondless in G. Both bondless and bondlessly surrounded molecules create difficulties in applying the rainbow matching lemma (Lemma 2.5.4), which in turn is used to assign absorption structures to vertices. Therefore, it will become important that we bound the number of each, and show that they are well spread out.

Lemma 2.7.3. *Let $\varepsilon > 0$ and $\ell, s, n \in \mathbb{N}$ be such that $s < n$, $\ell \leq n - s$ and $1/\ell \ll \varepsilon$. Then, for any (s,ℓ)-molecule $\mathcal{M} \subseteq \mathcal{Q}^n$, the probability that it is bondless in $\mathcal{Q}^n_\varepsilon$ is at most $2^{s+1-\varepsilon 2^\ell/4}$.*

Proof. Fix an (s,ℓ)-molecule $\mathcal{M} = \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_{2^s} \subseteq \mathcal{Q}^n$. Consider a pair of consecutive atoms $\mathcal{A}_i, \mathcal{A}_{i+1} \subseteq \mathcal{M}$, for some $i \in [2^s]$. Let X_i be the number of edges between $\mathcal{A}_i$ and $\mathcal{A}_{i+1}$ in $\mathcal{Q}^n_\varepsilon$ whose endpoint in $\mathcal{A}_i$ is odd, and let Y_i be the number of such edges whose endpoint in $\mathcal{A}_i$ is even. We have that $X_i, Y_i \sim \mathrm{Bin}(2^{\ell-1}, \varepsilon)$. By Lemma 2.4.2, it follows that

$$\mathbb{P}[X_i < 100] \leq 2^{-\varepsilon 2^\ell/4},$$

and the same bound holds for $\mathbb{P}[Y_i < 100]$. By a union bound over all $i \in [2^s]$, we conclude that

$$\mathbb{P}[\mathcal{M} \text{ is bondless in } \mathcal{Q}^n_\varepsilon] \leq 2^{s+1-\varepsilon 2^\ell/4}. \qquad \square$$

Lemma 2.7.4. *Let $\varepsilon \in (0,1)$ and $\ell, n \in \mathbb{N}$ with $0 < 1/n \ll 1/\ell \ll \varepsilon$, and let $s := 10\ell$. Let $\mathfrak{M}$ be a collection of vertex-disjoint (s,ℓ)-molecules $\mathcal{M} \subseteq \mathcal{Q}^n$. For each $x \in V(\mathcal{Q}^n)$, let $N^{\mathfrak{M}}(x) := \{\mathcal{M} \in \mathfrak{M} : \mathrm{dist}(x,\mathcal{M}) = 1\}$. Assume that the following holds for every $x \in V(\mathcal{Q}^n)$:*

(BS) *for any direction $\hat{e} \in \mathcal{D}(\mathcal{Q}^n)$, there are at most $\sqrt{n}$ molecules $\mathcal{M} \in N^{\mathfrak{M}}(x)$ such that $\hat{e} \in \mathcal{D}(\mathcal{A})$ for all atoms $\mathcal{A} \in \mathcal{M}$.*

Then, with probability at least $1 - 2^{-n^{9/8}}$, for every $x \in V(\mathcal{Q}^n)$ we have that $B^{\ell^2}_{\mathcal{Q}^n}(x)$ intersects at most $n^{1/3}$ molecules from $\mathfrak{M}$ which are bondlessly surrounded in $\mathcal{Q}^n_\varepsilon$.

Proof. We begin by fixing an arbitrary vertex $x \in V(\mathcal{Q}^n)$ and an arbitrary set $B \subseteq \mathfrak{M}$ of $n^{1/3}$ molecules which intersect $B^{\ell^2}_{\mathcal{Q}^n}(x)$. We will estimate the probability that all of the molecules in B are bondlessly surrounded in $\mathcal{Q}^n_\varepsilon$, by considering the neighbourhoods of the different vertices which make up these molecules. If the probability of being bondlessly surrounded was independent over different molecules and vertices, then this would be a straightforward calculation. However, there are dependencies which we must consider: namely, when two different molecules have edges to the same third molecule. We will first bound the number of such configurations in $\mathcal{Q}^n$. Since the molecules in $\mathfrak{M} \supseteq B$ are vertex-disjoint, it follows that, if two of these molecules are adjacent in $\mathcal{Q}^n$, then all of their atoms are pairwise adjacent in each of the layers, via clones of the same edges. Thus, we can restrict the analysis to a single layer.

Fix a layer L and let $\mathfrak{A}$ be the collection of atoms obtained by intersecting each molecule $\mathcal{M} \in \mathfrak{M}$ with L. Let $\mathfrak{A}_B \subseteq \mathfrak{A}$ be the set of such atoms whose molecules lie in B. Fix an atom $\mathcal{A} \in \mathfrak{A}_B$, and let $y \in V(\mathcal{A})$ be a fixed vertex. We say an atom $\mathcal{A}' \in \mathfrak{A}$ is *y-dependent* if there exists $\mathcal{A}'' \in \mathfrak{A}_B$, $\mathcal{A}'' \neq \mathcal{A}$, such that $\mathrm{dist}(y,\mathcal{A}') = \mathrm{dist}(\mathcal{A}',\mathcal{A}'') = 1$. The following claim will allow us to bound the number of y-dependent atoms.

Claim 2.7.1. *Fix $\mathcal{A}'' \in \mathfrak{A}_B$ with $\mathcal{A}'' \neq \mathcal{A}$. Then, the number of $\mathcal{A}' \in \mathfrak{A}$ for which $\mathrm{dist}(y,\mathcal{A}') = \mathrm{dist}(\mathcal{A}',\mathcal{A}'') = 1$ is at most $2^\ell(2+\sqrt{n})$.*

Proof. Let $z \in V(\mathcal{A}'')$ and let $\hat{e} \in \mathcal{D}(y,z)$. Let $\mathcal{A}' \in \mathfrak{A}$ be such that $\mathrm{dist}(y,\mathcal{A}') = \mathrm{dist}(z,\mathcal{A}') = 1$. Suppose first that $\hat{e} \notin \mathcal{D}(\mathcal{A}')$. Then we must have either $y+\hat{e} \in V(\mathcal{A}')$ or $z+\hat{e} \in V(\mathcal{A}')$. Since all the atoms in $\mathfrak{A}$ are vertex-disjoint, this leaves only two possibilities for $\mathcal{A}'$. Alternatively, suppose $\hat{e} \in \mathcal{D}(\mathcal{A}')$. Then, by (BS) applied with y playing the role of x, we have at most $\sqrt{n}$ possibilities for $\mathcal{A}'$. Finally, by considering all $z \in V(\mathcal{A}'')$ we prove the claim. ◀

By considering all possibilities for $\mathcal{A}'' \in \mathfrak{A}_B$, since $|\mathfrak{A}_B| = n^{1/3}$, it follows by Claim 2.7.1 that the number of y-dependent atoms is at most $n^{6/7}$. For each $y \in V(\mathcal{A})$, let $N'(y) \subseteq N^{\mathfrak{M}}(y)$ be given by removing from $N^{\mathfrak{M}}(y)$ all molecules which contain a y-dependent atom. It follows that $|N'(y)| = |N^{\mathfrak{M}}(y)| - o(n)$ for every $y \in V(\mathcal{A})$.

Let $\mathcal{M}_{\mathcal{A}} \in \mathfrak{M}$ be the molecule containing $\mathcal{A}$. For each vertex $y \in V(\mathcal{A})$, let $\mathcal{E}_y$ be the event that $N'(y)$ contains at least $n/2^{\ell+5s+1}$ molecules $\mathcal{M} \in \mathfrak{M}$ which are bondless in $\mathcal{Q}^n_\varepsilon$. Then, $|N'(y)| \geq n/2^{\ell+5s+2}$. Moreover, we only consider here those vertices $y \in V(\mathcal{A})$ for which $|N^{\mathfrak{M}}(y)| \geq n/2^{\ell+5s+1}$, since otherwise y cannot contribute towards $\mathcal{M}_{\mathcal{A}}$ being bondlessly surrounded. Fix such a vertex y. Let Y be the number of atoms $\mathcal{A} \in N'(y)$ which correspond to molecules which are bondless in $\mathcal{Q}^n_\varepsilon$. Note that Y is a sum of independent indicator variables. By Lemma 2.7.3, we have that $\mathbb{E}[Y] \leq 2^{s+1-\varepsilon 2^\ell/4} n$. In order to derive a lower bound for $\mathbb{E}[Y]$, note that the probability that an (s,ℓ)-molecule $\mathcal{M}$ is bondless can be bounded from below by the probability that there are no edges between two fixed consecutive atoms $\mathcal{A}_1, \mathcal{A}_2 \subseteq \mathcal{M}$, whose endpoints in $\mathcal{A}_1$ are even. This occurs with probability $(1-\varepsilon)^{2^{\ell-1}}$. Thus,

$$\mathbb{E}[Y] \geq (1-\varepsilon)^{2^{\ell-1}} |N'(y)| \geq (1-\varepsilon)^{2^{\ell-1}} (n/2^{\ell+5s+2}).$$

By Lemma 2.4.4, we have that $\mathbb{P}[\mathcal{E}_y] \leq 2^{-cn}$, for some constant $c > 0$ which depends on ℓ and ε. For each atom $\mathcal{A} \in \mathfrak{A}_B$, let $\mathcal{B}_{\mathcal{A}}$ be the event that there exists a vertex $y \in V(\mathcal{A})$ such that $\mathcal{E}_y$ holds. Let $\mathcal{B} \coloneqq \bigwedge_{\mathcal{A} \in \mathfrak{A}_B} \mathcal{B}_{\mathcal{A}}$. Note that the definition of $N'(y)$ ensures that the events

$\mathcal{B}_{\mathcal{A}}$ with $\mathcal{A} \in \mathfrak{A}_B$ are pairwise independent. Thus,

$$\mathbb{P}[\mathcal{E}] \leq (2^{\ell - cn})^{n^{1/3}} < 2^{-n^{5/4}}.$$

In turn, this means that the probability that all molecules $\mathcal{M} \in B$ are bondlessly surrounded is bounded from above by $2^{-n^{5/4}}$. Lemma 2.7.4 now follows by a union bound over the 2^n choices for x and the at most $\binom{n^{\ell^2}}{n^{1/3}}$ choices for B. □

Finally, we will show that 'scant' molecules are not too clustered. (We will later define a vertex molecule as 'scant' –with respect to a graph H and a reservoir R– if one of its vertices v_i has the property that few of its neighbours lie in the i-th clone of R.)

Lemma 2.7.5. *Let $C, s, n \in \mathbb{N}$ such that $0 < 1/n \ll 1/C \ll \alpha, \delta \leq 1$ and $1/n \ll 1/s$. Let $H \subseteq \mathcal{Q}^n$ be such that $\delta(H) \geq \alpha n$. For each $v \in V(\mathcal{Q}^{n-s})$ and each $i \in [2^s]$, let v_i be the i-th clone of v, and let $\mathcal{M}_v \coloneqq \{v_i : i \in [2^s]\}$. Let $R \sim Res(\mathcal{Q}^{n-s}, \delta)$ and, for each $i \in [2^s]$, let R_i be the i-th clone of R. Let*

$$B \coloneqq \{\mathcal{M}_v \mid v \in V(\mathcal{Q}^{n-s}),\ \text{there exists } i \in [2^s] : e_H(v_i, R_i) < \alpha\delta n/4\}.$$

Let $\mathcal{E}$ be the event that there exists some $u \in V(\mathcal{Q}^{n-s})$ such that $B^{10\ell}_{\mathcal{Q}^{n-s}}(u)$ contains more than C vertices $v \in V(\mathcal{Q}^{n-s})$ with $\mathcal{M}_v \in B$. Then, $\mathbb{P}[\mathcal{E}] < e^{-n}$.

Proof. Let $u \in V(\mathcal{Q}^{n-s})$ and let $D \subseteq B^{10\ell}_{\mathcal{Q}^{n-s}}(u)$ be a set of C vertices. Let $D' \coloneqq \bigcup_{x,y \in D : x \neq y} N_{\mathcal{Q}^{n-s}}(x) \cap N_{\mathcal{Q}^{n-s}}(y)$. Since any pair of distinct vertices share at most two neighbours, we have that $|D'| \leq 2\binom{C}{2}$. For each $i \in [2^s]$, we denote the i-th clone of D' by D'_i, and let $R'_i \coloneqq R_i \setminus D'_i$.

For each $x \in V(\mathcal{Q}^n)$, let $i(x)$ be the unique index $i \in [2^s]$ such that $x \in V(L_i)$. Observe that $e_H(x, V(L_{i(x)})) > 2\alpha n/3$ for every $x \in V(\mathcal{Q}^n)$. For each $x \in V(\mathcal{Q}^n)$, let $\mathcal{E}_x$ be the event that $e_H(x, R_{i(x)}) \leq \alpha\delta n/4$, and let $\mathcal{E}'_x$ be the event that $e_H(x, R'_{i(x)}) \leq \alpha\delta n/4$. It follows by Lemma 2.4.2 that $\mathbb{P}[\mathcal{E}'_x] \leq e^{-\alpha\delta n/16}$ for all $x \in V(\mathcal{Q}^n)$. For each $v \in V(\mathcal{Q}^{n-s})$, let $\mathcal{E}_v$ and $\mathcal{E}'_v$ be the events that there exists $i \in [2^s]$ such that $\mathcal{E}_{v_i}$ and $\mathcal{E}'_{v_i}$ hold, respectively. By a union bound, it follows that $\mathbb{P}[\mathcal{E}'_v] \leq 2^s e^{-\alpha\delta n/16}$ for all $v \in V(\mathcal{Q}^{n-s})$. Finally, let $\mathcal{E}_D$ and $\mathcal{E}'_D$ be the events that $\mathcal{E}_v$ and $\mathcal{E}'_v$, respectively, hold for every $v \in D$. Note that the events in the collection $\{\mathcal{E}'_v : v \in V(\mathcal{Q}^{n-s})\}$ are mutually independent. Furthermore, since the event $\mathcal{E}_x$ implies $\mathcal{E}'_x$ for all $x \in V(\mathcal{Q}^n)$, we have that

$$\mathbb{P}[\mathcal{E}_D] \leq \mathbb{P}[\mathcal{E}'_D] \leq (2^s e^{-\alpha\delta n/16})^C < e^{-5n}.$$

Taking a union bound over all vertices u and over all choices of D we obtain the result. □

2.7.3 Connecting cubes

The hypercube satisfies some robust connectivity properties. The problem of (almost) covering $\mathcal{Q}^n$ with disjoint paths has been extensively studied.

In order to create a long cycle, which can be used to absorb all remaining vertices, while preserving the absorbing structure, we will make use of the robust connectivity properties of the hypercube. In particular, we will need several results which guarantee that, given any prescribed pairs of vertices in a slice, there is a spanning collection of vertex-disjoint paths, each of which uses the vertices of one of the given pairs as endpoints. We will also need similar results for almost spanning collections of paths, where these paths avoid a given prescribed vertex. Throughout this subsection we denote by uv the edge between two given adjacent vertices u and v (instead of $\{u, v\}$).

The following lemma will be essential for us. It follows from some results of Dvořák and Gregor [39, Corollary 5.2].

Lemma 2.7.6. *For all $n \geq 100$, the graph $\mathcal{Q}^n$ satisfies the following.*

(i) *Let $m \in [25]$ and let $\{u_i, v_i\}_{i\in[m]}$ be disjoint pairs of vertices with $u_i \neq_{\mathrm{p}} v_i$ for all $i \in [m]$. Then, there exist m vertex-disjoint paths $\mathcal{P}_1, \ldots, \mathcal{P}_m \subseteq \mathcal{Q}^n$ such that, for each $i \in [m]$, $\mathcal{P}_i$ is a (u_i, v_i)-path, and $\bigcup_{i\in[m]} V(\mathcal{P}_i) = V(\mathcal{Q}^n)$.*

(ii) *Let* $x \in V(\mathcal{Q}^n)$*. Let* $m \in [25]$ *and let* $\{u_i, v_i\}_{i\in[m]}$ *be disjoint pairs of vertices of* $\mathcal{Q}^n - \{x\}$ *such that* $u_1, v_1 \neq_{\mathrm{p}} x$ *and* $u_i \neq_{\mathrm{p}} v_i$ *for all* $i \in [m] \setminus \{1\}$*. Then, there exist* m *vertex-disjoint paths* $\mathcal{P}_1, \ldots, \mathcal{P}_m \subseteq \mathcal{Q}^n$ *such that, for each* $i \in [m]$*,* $\mathcal{P}_i$ *is a* (u_i, v_i)*-path, and* $\bigcup_{i\in[m]} V(\mathcal{P}_i) = V(\mathcal{Q}^n) \setminus \{x\}$*.*

(iii) *Let* $\{u_i, v_i\}_{i\in[2]}$ *be disjoint pairs of vertices with* $u_i =_{\mathrm{p}} v_i$ *for all* $i \in [2]$ *and* $u_1 \neq_{\mathrm{p}} u_2$*. Then, there exist two vertex-disjoint paths* $\mathcal{P}_1, \mathcal{P}_2 \subseteq \mathcal{Q}^n$ *such that, for each* $i \in [2]$*,* $\mathcal{P}_i$ *is a* (u_i, v_i)*-path, and* $V(\mathcal{P}_1) \cup V(\mathcal{P}_2) = V(\mathcal{Q}^n)$*.*

We now motivate the statement (as well as the proof) of Lemma 2.7.8, which is the main result of this subsection. We are given a slice $\mathcal{M}^*$ of a molecule $\mathcal{M} \subseteq \mathcal{Q}^n$ which is bonded in a graph $G \subseteq \mathcal{Q}^n$. Furthermore, we are given collections of vertices L, R (which are part of absorbing cube structures), and S (which, when constructing a long cycle, will be used to enter and leave $\mathcal{M}^*$). More specifically, we have that

- L will have size 0 or 2, and will consist of left absorber tips. If it has size 2, the vertices will have opposite parities. These must be avoided by our connecting paths, so that we can make use of the absorbing structures we have put in place (see the discussion in Section 2.2).
- R will consist of the pairs of right absorber tip and third absorber vertex. These must be connected via an edge with the paths we find.
- S will consist of a set of pairs of vertices $\{u, v\}$ with $u \neq_{\mathrm{p}} v$. Later, when creating a long cycle, u will be a vertex through which we enter $\mathcal{M}^*$ from a different molecule, and v will be the next vertex from which we leave $\mathcal{M}^*$ (with respect to some ordering). Each of our paths will be a (u, v)-path, for some such pair $\{u, v\}$.

In order to find our paths, we will call on Lemma 2.7.6. To illustrate this, suppose $\mathcal{M}^*$ consists of the atoms $\mathcal{A}_1, \ldots, \mathcal{A}_t$, for some $t \in \mathbb{N}$. Suppose that $S = \{u, v\}$ with $u \in V(\mathcal{A}_1)$ and $v \in V(\mathcal{A}_t)$. Furthermore, suppose that $L, R = \varnothing$. To construct a path from u to v, we will first specify the edges used to pass between different atoms. For all $k \in [t-1]$, we choose an edge $v_k^{\uparrow}u_{k+1}^{\uparrow}$ from $\mathcal{A}_k$ to $\mathcal{A}_{k+1}$, thus $v_k^{\uparrow} \neq_{\mathrm{p}} u_{k+1}^{\uparrow}$. For technical reasons, we aim to have all the vertices $u_{k+1}^{\uparrow}$ of the same parity as u. We can then apply Lemma 2.7.6 to find a path from $u_{k+1}^{\uparrow}$ to $v_{k+1}^{\uparrow}$ which covers all of $V(\mathcal{A}_{k+1})$. Together with the edges $v_k^{\uparrow}u_{k+1}^{\uparrow}$, all these paths will form a single path from u to v which spans $V(\mathcal{M}^*)$. In the more general setting where $u \in V(\mathcal{A}_i)$ and $v \in V(\mathcal{A}_j)$ with $1 < i < j < t$, the (u, v)-path we construct would first pass down to $\mathcal{A}_1$, then up to $\mathcal{A}_t$ and, finally, back down to $\mathcal{A}_j$.

When $L \neq \varnothing$, due to vertex parities, the following issue can arise. Suppose $L = \{x, y\}$ with $x \in V(\mathcal{A}_1)$, $u \in V(\mathcal{A}_2)$, $y \in V(\mathcal{A}_3)$ and $v \in V(\mathcal{A}_j)$ for some $j > 3$ (and $R = \varnothing$). Furthermore, suppose that both u and x have odd parity. In line with the above description, the vertex $u_1^{\downarrow}$, through which we enter $\mathcal{A}_1$, would have odd parity. It follows that, since x also has odd parity, we cannot hope to construct a path which starts at $u_1^{\downarrow}$ and covers all of $V(\mathcal{A}_1) \setminus \{x\}$. The solution will be instead to pass up to $\mathcal{A}_3$ first (and, in general, to whichever atom contains y). Recall that, since x has odd parity, y must have even parity. We specify a vertex $u_3^{\uparrow}$ of odd parity, through which we enter $\mathcal{A}_3$, but then also specify a vertex $v_3^{\downarrow}$ of odd parity from which we will leave $\mathcal{A}_3$ to reenter $\mathcal{A}_2$. We now arrive back in $\mathcal{A}_2$ with a vertex $u_2^{\downarrow}$ of even parity. We will specify another vertex $v_2^{\downarrow}$ of odd parity from which we leave $\mathcal{A}_2$ and a vertex $u_1^{\downarrow}$ of even parity through which we enter $\mathcal{A}_1$. In this way, we can now apply Lemma 2.7.6 to find a path which starts at $u_1^{\downarrow}$ and covers all of $V(\mathcal{A}_1) \setminus \{x\}$, and which can be extended into a path from u to v covering all of $V(\mathcal{M}^*) \setminus L$.

There are several other instances which must be dealt with in a similar way. This is formalised by Lemma 2.7.8. Before proving this lemma, however, we need the following definition.

Definition 2.7.7 ((u, j, F, R)-alternating parity sequence)**.** *Let $\ell, s, t, n \in \mathbb{N}$ with $t \leq 2^s$ and $2 \leq \ell \leq n - s$. Let $G \subseteq \mathcal{Q}^n$. Let $\mathcal{M} = \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_{2^s} \subseteq \mathcal{Q}^n$ be an (s, ℓ)-molecule and let $\mathcal{M}^* = \mathcal{A}_{a+1} \cup \cdots \cup \mathcal{A}_{a+t}$, for some $a \in [2^s]$, be a slice of $\mathcal{M}$. Let $u \in V(\mathcal{A}_i)$, for* *some $i \in [a+t] \setminus [a]$. Let $j \in [a+t] \setminus [a]$, and let $F, R \subseteq V(\mathcal{M}^*)$. Suppose $i \le j$. Let $I_R \coloneqq \{k \in [j-i]_0 : |R \cap V(\mathcal{A}_{i+k})| \ge 1\}$. Assume that the following properties hold:*

- *For all $k \in [j-i]_0$ we have that $|R \cap V(\mathcal{A}_{i+k})| \in \{0, 2\}$.*
- *For each $k \in I_R$, the vertices in $R \cap V(\mathcal{A}_{i+k})$ are adjacent in $\mathcal{Q}^n$, and we write $R \cap V(\mathcal{A}_{i+k}) = \{w_k, z_k\}$ so that $w_k \neq_{\mathrm{p}} u$.*

Let $\mathcal{S}' = (u_0, v_1, u_1, \dots, v_{j-i}, u_{j-i})$ be a sequence of vertices satisfying the following properties:

(P0) *If $u \in R$, then $u_0 \coloneqq w_0$; otherwise, $u_0 \coloneqq u$.*

(P1) *For each $k \in [j-i]$ we have that $u_k =_{\mathrm{p}} u$.*

(P2) *For each $k \in [j-i]$ we have that $v_k \in V(\mathcal{A}_{i+k-1})$, $u_k \in V(\mathcal{A}_{i+k})$ and $v_k u_k \in E(G)$.*

(P3) *The vertices of $\mathcal{S}'$ other than u_0 avoid $F \cup R$.*

A (u, j, F, R)-alternating parity sequence *$\mathcal{S}$ in G is a sequence obtained from any sequence $\mathcal{S}'$ which satisfies (P0)–(P3) as follows. For each $k \in I_R \cap [j-i]$, replace each segment (v_k, u_k) of $\mathcal{S}'$ by (v_k, u_k, w_k, z_k).*

The case $i > j$ is defined similarly by replacing each occurrence of $[j-i]$ and $[j-i]_0$ in the above by $[i-j]$ and $[i-j]_0$, and each occurrence of $\mathcal{A}_{i+k}$ and $\mathcal{A}_{i+k-1}$ by $\mathcal{A}_{i-k}$ and $\mathcal{A}_{i-k+1}$.

Given an alternating parity sequence $\mathcal{S}$, we will denote by $\mathcal{S}^-$ the sequence obtained from $\mathcal{S}$ by deleting its initial element.

Lemma 2.7.8. *Let $n, s, \ell \in \mathbb{N}$ be such that $s \ge 4$ and $100 \le \ell \le n - s$. Let $G \subseteq \mathcal{Q}^n$ and consider any (s, ℓ)-molecule $\mathcal{M} = \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_{2^s} \subseteq \mathcal{Q}^n$ which is bonded in G. Let $\mathcal{M}^* = \mathcal{A}_{a+1} \cup \cdots \cup \mathcal{A}_{a+t}$, for some $a \in [2^s]$ and $t \ge 10$, be a slice of $\mathcal{M}$. Moreover, consider the following sets.*

(C1) *Let $L \subseteq V(\mathcal{M}^*)$ be a set of size $|L| \in \{0, 2\}$ such that, if $L = \{x, y\}$, then $x \in V(\mathcal{A}_i)$ and $y \in V(\mathcal{A}_j)$ with $i \neq j$ and $x \neq_{\mathrm{p}} y$.*

(C2) *Let $R \subseteq V(\mathcal{M}^*) \setminus L$ be a (possibly empty) set of vertices with $|R| \leq 10$ such that, for all $k \in [a+t] \setminus [a]$, we have $|R \cap V(\mathcal{A}_k)| \in \{0, 2\}$ and, if $|R \cap V(\mathcal{A}_k)| = 2$, then $R \cap V(\mathcal{A}_k) = \{w_k, z_k\}$ satisfies that $w_k z_k \in E(\mathcal{M}^*)$ and, if $|L| = 2$, then $k \notin \{i, j\}$.*

(C3) *Let $m \in [14]$ and consider m vertex-disjoint pairs $\{u_r, v_r\}_{r \in [m]}$, where $u_r, v_r \in V(\mathcal{M}^*) \setminus L$ and $u_r \neq_{\mathrm{p}} v_r$ for all $r \in [m]$, such that, for each $r \in [m]$, we have $u_r \in V(\mathcal{A}_{i_r})$ and $v_r \in V(\mathcal{A}_{j_r})$. Assume, furthermore, that for each $t' \in [t]$ we have that $|\bigcup_{r \in [m]} \{u_r, v_r\} \cap V(\mathcal{A}_{a+t'}) \cap R| \leq 1$.*

Then, there exist vertex-disjoint paths $\mathcal{P}_1, \ldots, \mathcal{P}_m \subseteq \mathcal{M}^* \cup G$ such that, for each $r \in [m]$, $\mathcal{P}_r$ is a (u_r, v_r)-path, $\bigcup_{r \in [m]} V(\mathcal{P}_r) = V(\mathcal{M}^*) \setminus L$, and every pair $\{w_k, z_k\}$ with $k \in [a+t] \setminus [a]$ is an edge of some $\mathcal{P}_r$.

Proof. By relabelling the atoms, we may assume that $\mathcal{M}^* = \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_t$. Let $S \coloneqq \{u_r, v_r : r \in [m]\}$. By relabelling the vertices, we may assume that $i_r \leq j_r$ for all $r \in [m]$ and (if $L \neq \varnothing$) $i < j$. Let $I_L \coloneqq \{k \in [t] : L \cap V(\mathcal{A}_k) \neq \varnothing\}$, $I_R \coloneqq \{k \in [t] : R \cap V(\mathcal{A}_k) \cap S \neq \varnothing\}$ and $R^* \coloneqq R \setminus \bigcup_{k \in I_R} V(\mathcal{A}_k)$. Note that $I_L = \varnothing$ or $I_L = \{i, j\}$ and $I_L \cap I_R = \varnothing$. For each $r \in [m]$, let $I_R^r \coloneqq \{k \in \{i_r, j_r\} : R \cap V(\mathcal{A}_k) \cap \{u_r, v_r\} \neq \varnothing\}$, so that $I_R = \bigcup_{r=1}^m I_R^r$. Without loss of generality, we may also assume that, for each $r \in [m]$, if $u_r \in R$, then $u_r = z_{i_r}$, and if $v_r \in R$, then $v_r = w_{j_r}$. Similarly, for each $k \in [t] \setminus I_R$, if $R \cap V(\mathcal{A}_k) = \{w_k, z_k\}$, we may assume that $w_k \neq_{\mathrm{p}} u_1$.

For each $r \in [m]$, we will create a list $\mathcal{L}_r$ of vertices. We will refer to $\mathcal{L}_r$ as the *skeleton* for $\mathcal{P}_r$. We will later use these skeletons to construct the vertex-disjoint paths via Lemma 2.7.6. For each $r \in [m]$, we will write L_r^* for the (unordered) set of vertices in $\mathcal{L}_r$. In order to construct each $\mathcal{L}_r$, we will start with an empty list and update it in (possibly) several steps, by concatenating alternating parity sequences. Whenever $\mathcal{L}_r$ is updated, we implicitly update L_r^*. In the end, for each $r \in [m]$ we will have a list of vertices $\mathcal{L}_r = (x_1^r, \ldots, x_{\ell_r}^r)$. For each $r \in [m]$ and $k \in [t]$, let $I_r(k) \coloneqq \{h \in [\ell_r - 1] : 2 \nmid h \text{ and } x_h^r, x_{h+1}^r \in V(\mathcal{A}_k)\}$. We will require the $\mathcal{L}_r$ to be pairwise vertex-disjoint. Furthermore, we will require that they satisfy the following properties:

($\mathcal{L}1$) For all $r \in [m]$ we have that ℓ_r is even.

($\mathcal{L}2$) For all $r \in [m]$ and $h \in [\ell_r - 1]$, if h is odd, then $x^r_h, x^r_{h+1} \in V(\mathcal{A}_k)$, for some $k \in [t]$; if h is even, then $x^r_h x^r_{h+1} \in E(G \cup \mathcal{M}^*)$.

($\mathcal{L}3$) For all $k \in [t]$ we have that $1 \leq |I_1(k)| \leq 6$ and $|I_r(k)| \leq 1$ for all $r \in [m] \setminus \{1\}$.

$(\mathcal{L}4)_1$ For each $k \in [t] \setminus (I_L \cup I^1_R)$ and each $h \in I_1(k)$, we have $x^1_h \neq_{\mathrm{p}} x^1_{h+1}$. For each $k \in I_L \cup I^1_R$, for all but one $h \in I_1(k)$ we have $x^1_h \neq_{\mathrm{p}} x^1_{h+1}$, while for the remaining index $h \in I_1(k)$ we have that $x^1_h =_{\mathrm{p}} x^1_{h+1}$ and their parity is opposite to that of the unique vertex in $L \cap V(\mathcal{A}_k)$ if $k \in I_L$ and to that of the unique vertex in $\{w_k, z_k\} \cap \{u_1, v_1\}$ if $k \in I^1_R$.

$(\mathcal{L}4)_r$ For each $r \in [m] \setminus \{1\}$, the following holds. For each $k \in [t] \setminus I^r_R$ and each $h \in I_r(k)$, we have $x^r_h \neq_{\mathrm{p}} x^r_{h+1}$. For each $k \in I^r_R$, for all but one $h \in I_r(k)$ we have $x^r_h \neq_{\mathrm{p}} x^r_{h+1}$, while for the remaining index $h \in I_r(k)$ we have that $x^r_h =_{\mathrm{p}} x^r_{h+1}$ and their parity is opposite to that of the unique vertex in $\{w_k, z_k\} \cap \{u_r, v_r\}$.

($\mathcal{L}5$) For each $r \in [m]$, we have the following. If $u_r \notin R$, then $u_r = x^r_1$. If $v_r \notin R$, then $v_r = x^r_{\ell_r}$. If $u_r \in R$ (and thus $u_r = z_{i_r}$), then $w_{i_r} = x^r_1$ and $u_r \notin L^*_1 \cup \cdots \cup L^*_m$. If $v_r \in R$ (and thus $v_r = w_{j_r}$), then $z_{j_r} = x^r_{\ell_r}$ and $v_r \notin L^*_1 \cup \cdots \cup L^*_m$.

($\mathcal{L}6$) Every pair (w_k, z_k) with $\{w_k, z_k\} \subseteq R^*$ is contained in $\mathcal{L}_1$ and z_k directly succeeds w_k.

We begin by constructing $\mathcal{L}_1$. Let $\mathcal{L}_1 := \varnothing$ and let $F := L \cup R \cup S$. If $i_1 = 1$ and $R^* \cap V(\mathcal{A}_1) = \{w_1, z_1\}$, then let $\mathcal{S}_1 := (u_1, w_1, z_1)$. If $i_1 = 1$ and $u_1 \in R$, then let $\mathcal{S}_1 := (u_1)$. Otherwise, let $\mathcal{S}_1$ be a $(u_1, 1, F, (R \cap V(\mathcal{A}_{i_1})) \cup (R^* \cap V(\mathcal{A}_1)))$-alternating parity sequence. Let $\mathcal{L}_1 := \mathcal{S}_1$. Note that the existence of such a sequence $\mathcal{S}_1$ is guaranteed by our assumption that $\mathcal{M}$ is bonded in G. To see this, note that all edges of G required by $\mathcal{S}_1$ (that is, the pairs $\{v_k, u_k\}$ in Definition 2.7.7) need to be chosen so that they do not have an endpoint in F; given any particular pair of consecutive atoms, this forbids at most 30 edges between these two atoms (26 because of S and 4 because of $L \cup R$).

We will now update $\mathcal{L}_1$. While doing so, we will update F and consider several alternating parity sequences. The existence of each of these follows a similar argument to the above. For any given pair of consecutive atoms, every time we update F, the set of forbidden edges will increase its size by at most 3. We will update F at most four times, so F will forbid at most 42 edges between any pair of consecutive atoms. Thus, by the definition of bondedness, each of the alternating parity sequences required below actually exists.

Let $u_1^{\downarrow}$ be the last vertex in $\mathcal{L}_1$. Note that $u_1^{\downarrow} =_{\mathrm{p}} u_1$ by Definition 2.7.7(P1). We update F as $F := F \cup L_1^*$. For the next step in the construction of $\mathcal{L}_1$, there are three cases to consider, depending on the size of L and, if $|L| = 2$, the relative parities of x and u_1. If $i_1 = 1$ and $u_1 \in R$, let $R^{\diamond} := R^* \cup \{w_1, z_1\}$; otherwise, let $R^{\diamond} := R^*$.

Case 1: $L = \varnothing$.

Let $\mathcal{S}_2$ be a $(u_1^{\downarrow}, t, F, R^{\diamond})$-alternating parity sequence. If $i_1 = 1$ and $u_1 \in R$, update $\mathcal{L}_1$ as $\mathcal{L}_1 := \mathcal{S}_2$. Otherwise, update $\mathcal{L}_1$ as $\mathcal{L}_1 := \mathcal{L}_1\mathcal{S}_2^-$. Update $F := F \cup L_1^*$.

Case 2: $|L| = 2$ and $x \neq_{\mathrm{p}} u_1$.

Let $\mathcal{S}_2$ be a $(u_1^{\downarrow}, i, F, R^{\diamond})$-alternating parity sequence. If $i_1 = 1$ and $u_1 \in R$, update $\mathcal{L}_1$ as $\mathcal{L}_1 := \mathcal{S}_2$. Otherwise, update $\mathcal{L}_1 := \mathcal{L}_1\mathcal{S}_2^-$. Update $F := F \cup L_1^*$. Choose any vertex $u_i^* \in V(\mathcal{A}_i)$ with $u_i^* \neq_{\mathrm{p}} u_1$, and let $\mathcal{S}_3$ be a $(u_i^*, j, F, R^{\diamond})$-alternating parity sequence. Update $\mathcal{L}_1 := \mathcal{L}_1\mathcal{S}_3^-$ and $F := F \cup L_1^*$. Let v^- be the final vertex of $\mathcal{S}_2$, and let v^+ be the second vertex of $\mathcal{S}_3$. Note that v^- and v^+ appear consecutively in $\mathcal{L}_1$ and that $v^- =_{\mathrm{p}} v^+ =_{\mathrm{p}} u_1 \neq_{\mathrm{p}} x$. Finally, choose any vertex $u_j^* \in V(\mathcal{A}_j)$ with $u_j^* =_{\mathrm{p}} u_1$, let $\mathcal{S}_4$ be a $(u_j^*, t, F, R^{\diamond})$-alternating parity sequence, and update $\mathcal{L}_1 := \mathcal{L}_1\mathcal{S}_4^-$ and $F := F \cup L_1^*$. Let w^- be the final vertex of $\mathcal{S}_3$, and let w^+ be the second vertex of $\mathcal{S}_4$. We then have that w^- and w^+ appear consecutively in $\mathcal{L}_1$, and $w^- =_{\mathrm{p}} w^+ \neq_{\mathrm{p}} y, u_1$. Moreover, the final vertex $u_t^{\uparrow}$ of $\mathcal{L}_1$ satisfies $u_t^{\uparrow} =_{\mathrm{p}} u_1$.

Case 3: $|L| = 2$ and $x =_{\mathrm{p}} u_1$.

Let $\mathcal{S}_2$ be a $(u_1^{\downarrow}, j, F, R^{\diamond})$-alternating parity sequence. If $i_1 = 1$ and $u_1 \in R$, update $\mathcal{L}_1$ as

$\mathcal{L}_1 \coloneqq \mathcal{S}_2$; otherwise, update $\mathcal{L}_1 \coloneqq \mathcal{L}_1\mathcal{S}_2^-$. Update $F \coloneqq F \cup L_1^*$. Next, let $u_j^* \in V(\mathcal{A}_j)$ be a vertex with $u_j^* \neq_{\mathrm{p}} u_1$ and let $\mathcal{S}_3$ be a $(u_j^*, i, F, \varnothing)$-alternating parity sequence. Update $\mathcal{L}_1 \coloneqq \mathcal{L}_1\mathcal{S}_3^-$ and $F \coloneqq F \cup L_1^*$. Finally, let $u_i^* \in V(\mathcal{A}_i)$ be a vertex with $u_i^* =_{\mathrm{p}} u_1$ and let $\mathcal{S}_4$ be a $(u_i^*, t, F, R^* \cap \bigcup_{k=j+1}^{t} V(\mathcal{A}_k))$-alternating parity sequence. Update $\mathcal{L}_1 \coloneqq \mathcal{L}_1\mathcal{S}_4^-$ and $F \coloneqq F \cup L_1^*$.

In each of the three cases, let $u_t^\uparrow$ denote the last vertex in $\mathcal{L}_1$. Note that, by Definition 2.7.7(P1), we have $u_t^\uparrow =_{\mathrm{p}} u_1$, and recall that $v_1 \neq_{\mathrm{p}} u_1$. Let $\mathcal{S}_5$ be a $(u_t^\uparrow, j_1, F, \varnothing)$-alternating parity sequence. Update $\mathcal{L}_1 \coloneqq \mathcal{L}_1\mathcal{S}_5^-$. Again by Definition 2.7.7(P1), we have that the final vertex u^* of $\mathcal{L}_1$ is such that $u^* =_{\mathrm{p}} u_t^\uparrow =_{\mathrm{p}} u_1 \neq_{\mathrm{p}} v_1$. Finally, if $v_1 \in R$, update $\mathcal{L}_1 \coloneqq \mathcal{L}_1(z_{j_1})$; otherwise, update it as $\mathcal{L}_1 \coloneqq \mathcal{L}_1(v_1)$. Observe that $\mathcal{L}_1$ satisfies $(\mathcal{L}1)$–$(\mathcal{L}3)$, $(\mathcal{L}4)_1$, $(\mathcal{L}5)$ and $(\mathcal{L}6)$ for the case $r = 1$ by construction.

We now construct $\mathcal{L}_r$ for all $r \in [m]\setminus\{1\}$. For each $r \in [m]\setminus\{1\}$, we proceed iteratively as follows. Let $\mathcal{L}_r \coloneqq \varnothing$ and $F_r \coloneqq L \cup R \cup S \cup \bigcup_{r' \in [r-1]} L_{r'}^*$. Let $\mathcal{S}^r$ be a $(u_r, j_r, F_r, R \cap V(\mathcal{A}_{i_r}))$-alternating parity sequence and update $\mathcal{L}_r$ as $\mathcal{L}_r \coloneqq \mathcal{S}^r$. If $v_r \in R$, update $\mathcal{L}_r \coloneqq \mathcal{L}_r(z_{j_r})$; otherwise, update $\mathcal{L}_r \coloneqq \mathcal{L}_r(v_r)$. Note that each sequence $\mathcal{S}^r$ requires the existence of at most one edge of G, which has to avoid F_r, between any pair of consecutive atoms of $\mathcal{M}^*$. In a similar way to what was discussed above, at most three choices of such edges can be forbidden every time we add a new alternating parity sequence to F. Since for each $r \in [m] \setminus \{1\}$ we consider one new sequence, by the time we consider F_m we have increased the number of forbidden edges by at most $3(m-1) \leq 39$. This gives a total of at most 81 forbidden edges and, thus, the existence of the sequences $\mathcal{S}^r$ is guaranteed by the assumption that $\mathcal{M}$ is bonded in G. Moreover, the lists $\mathcal{L}_1, \ldots, \mathcal{L}_r$ now satisfy $(\mathcal{L}1)$–$(\mathcal{L}6)$.

We are now in a position to apply Lemma 2.7.6. For each $k \in [t]$, let $t_k \coloneqq \sum_{r \in [m]} |I_r(k)|$. Furthermore, for any $r \in [m]$ and $k \in [t]$, for each $h \in I_r(k)$, we refer to the pair x_h^r, x_{h+1}^r as a *matchable pair*. By $(\mathcal{L}3)$, $(\mathcal{L}4)_1$, $(\mathcal{L}4)_r$ and Lemma 2.7.6(i), each atom $\mathcal{A}_k$ with $k \in [t]\setminus(I_L \cup I_R)$ can be covered by t_k vertex-disjoint paths, each of whose endpoints are a matchable pair contained in $\mathcal{A}_k$. Similarly, by $(\mathcal{L}3)$, $(\mathcal{L}4)_1$, $(\mathcal{L}4)_r$ and Lemma 2.7.6(ii), each atom $\mathcal{A}_k$ with

$k \in I_L \cup I_R$ contains t_k vertex-disjoint paths, each of whose endpoints are a matchable pair in $\mathcal{A}_k$ such that the union of these t_k paths covers precisely $V(\mathcal{A}_k) \setminus (L \cup (S \cap R))$. (Recall that by (C2) and (C3) the set $V(\mathcal{A}_k) \cap (L \cup (S \cap R))$ consists of a single vertex if $k \in I_L \cup I_R$.) For each matchable pair x_h^r, x_{h+1}^r in $\mathcal{A}_k$, let us denote the corresponding path by $\mathcal{P}_{x_h^r, x_{h+1}^r}$.

The paths $\mathcal{P}_1, \ldots, \mathcal{P}_m$ required for Lemma 2.7.8 can now be constructed as follows. For each $r \in [m]$, let $\mathcal{P}_r$ be the path obtained from the concatenation of the paths $\mathcal{P}_{x_h^r, x_{h+1}^r}$, for each odd $h \in [\ell_r]$, via the edges $x_h^r x_{h+1}^r$ for $h \in [\ell_r - 1]$ even. By ($\mathcal{L}5$), if $\mathcal{P}_r$ does not contain u_r, then $\mathcal{P}_r$ starts in w_{i_r}, and u_r does not lie in any other path; therefore, we can update $\mathcal{P}_r$ as $\mathcal{P}_r \coloneqq u_r \mathcal{P}_r$. Similarly, if $\mathcal{P}_r$ does not contain v_r, then $\mathcal{P}_r$ ends in z_{j_r} and v_r does not lie in any other path, and thus we can update $\mathcal{P}_r$ as $\mathcal{P}_r \coloneqq \mathcal{P}_r v_r$. It follows that $\bigcup_{r \in [m]} V(\mathcal{P}_r) = V(\mathcal{M}^*) \setminus L$, and thus the paths $\mathcal{P}_r$ are as required in Lemma 2.7.8. $\square$

We also need the following simpler result. Its proof follows similar ideas as those present in the proof of Lemma 2.7.8. For the sake of completeness, we include the proof of Lemma 2.7.9 in Appendix A of the arXiv version of this paper. We point out here that Lemma 2.7.6(iii) is only needed for this proof.

Lemma 2.7.9. *Let $n, s, \ell \in \mathbb{N}$ be such that $4 \leq s$ and $100 \leq \ell \leq n - s$. Let $G \subseteq \mathcal{Q}^n$ and consider any (s, ℓ)-molecule $\mathcal{M} = \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_{2^s} \subseteq \mathcal{Q}^n$ which is bonded in G. Let $\mathcal{M}^* = \mathcal{A}_{a+1} \cup \cdots \cup \mathcal{A}_{a+t}$, for some $a \in [2^s]$ and $t \geq 10$, be a slice of $\mathcal{M}$. Moreover, consider the following sets.*

(C′1) *Let $L \subseteq V(\mathcal{M}^*)$ be a set of size $|L| \in \{0, 2\}$ such that, if $L = \{x, y\}$, then $x \in V(\mathcal{A}_i)$ and $y \in V(\mathcal{A}_j)$, with $i \neq j$ and $x \neq_{\mathrm{p}} y$.*

(C′2) *Let $R \subseteq V(\mathcal{M}^*) \setminus L$ be a (possibly empty) set of vertices with $|R| \leq 10$ such that, for all $k \in [a + t] \setminus [a]$, we have $|R \cap V(\mathcal{A}_k)| \in \{0, 2\}$ and, if $|R \cap V(\mathcal{A}_k)| = 2$, then $R \cap V(\mathcal{A}_k) = \{w_k, z_k\}$ satisfies that $w_k z_k \in E(\mathcal{M}^*)$ and, if $|L| = 2$, then $k \notin \{i, j\}$.*

(C′3) *Consider two vertex-disjoint pairs* $\{u_r, v_r\}_{r\in[2]}$ *with* $u_1, u_2 \in V(\mathcal{A}_{a+1}) \setminus L$ *and* $v_1, v_2 \in V(\mathcal{A}_{a+t}) \setminus L$ *such that* $u_1 \neq_{\mathrm{p}} u_2$, $v_1 \neq_{\mathrm{p}} v_2$, $u_1 =_{\mathrm{p}} v_1$, *and* $|\{u_1, u_2\} \cap R|, |\{v_1, v_2\} \cap R| \leq 1$.

Then, there exist two vertex-disjoint paths $\mathcal{P}_1, \mathcal{P}_2 \subseteq \mathcal{M}^* \cup G$ *such that, for each* $r \in [2]$, $\mathcal{P}_r$ *is a* (u_r, v_r)*-path,* $V(\mathcal{P}_1) \cup V(\mathcal{P}_2) = V(\mathcal{M}^*) \setminus L$, *and every pair of the form* $\{w_k, z_k\} \subseteq R$ *with* $k \in [a+t] \setminus [a]$ *is an edge of either* $\mathcal{P}_1$ *or* $\mathcal{P}_2$.

2.7.4 Proof of Theorem 2.7.1

Proof of Theorem 2.7.1. Let $1/D, \delta' \ll 1$, and let

$$0 < 1/n_0 \ll \delta \ll 1/\ell \ll 1/k^*, \alpha' \ll \beta, 1/S' \ll 1/c, 1/D, \delta', \varepsilon, \alpha,$$

where $n_0, \ell, k^*, S', D \in \mathbb{N}$. Our proof assumes that n tends to infinity; in particular, $n \geq n_0$. Let $s := 10\ell$, $\Phi := 12\ell$ and and $\Psi := c\Phi$.

Observe that $\mathcal{Q}^n[\{0,1\}^s \times \{0\}^{n-s}] \cong \mathcal{Q}^s$ contains a Hamilton cycle. We fix an ordering of the layers $L_1, \ldots, L_{2^s}$ of $\mathcal{Q}^n$ induced by this Hamilton cycle (as defined in Section 2.7.1). If we view these layers as different subgraphs on the vertex set of $\mathcal{Q}^{n-s}$, we can define the *intersection graph* of the layers $I := \bigcap_{i=1}^{2^s} L_i$ (note that $I \cong \mathcal{Q}^{n-s}$) and, for any $G \subseteq \mathcal{Q}^n$, we denote $I(G) := \bigcap_{i=1}^{2^s} L_i(G)$. Note that, if $\mathcal{G} \subseteq I(G)$, then there is a clone of $\mathcal{G}$ in $L_i(G)$, for each $i \in [2^s]$. For each layer L, we denote by $\mathcal{G}_L$ the clone of $\mathcal{G}$ in $L(G)$. Observe that, for any $\eta \in [0,1]$, we have $I(\mathcal{Q}^n_\eta) \sim \mathcal{Q}^{n-s}_{\eta^{2^s}}$. We will sometimes write G_I for the subgraph of I where, for each $e \in E(I)$, we have $e \in E(G_I)$ whenever G contains some clone of e (thus, G_I is the 'union' of the subgraphs that G induces on each layer).

For each $i \in [7]$, let $\varepsilon_i := \varepsilon/7$ and let $G_i \sim \mathcal{Q}^n_{\varepsilon_i}$, where these graphs are taken independently. It is easy to see that $\bigcup_{i=1}^7 G_i \sim \mathcal{Q}^n_{\varepsilon'}$ for some $\varepsilon' < \varepsilon$. Thus, it suffices to show that a.a.s. there is a graph $G' \subseteq \bigcup_{i=1}^7 G_i$ with $\Delta(G') \leq \Phi$ such that, for every $F \subseteq \mathcal{Q}^n$ with $\Delta(F) \leq \Psi$, the graph $((H \cup \bigcup_{i=1}^7 G_i) \setminus F) \cup G'$ is Hamiltonian. In summary, we fix a near-spanning tree and reservoir in the intersection graph of G_1. G_6 is used to extend this tree to cover almost all of the vertices of the reservoir. G_7 is then used to alter the tree slightly at vertices which have few neighbours in the reservoir. We fix a near-spanning cube tiling of the intersection graph of G_3 and we use G_5 to fix good edge connectivity between different layers within molecules of the cubes of this tiling. Finally, G_2 and G_4 are used to find and fix suitable absorbing structures for each vertex of the hypercube. We now split our proof into several steps.

Step 1: Finding a tree and a reservoir. Consider the probability space $\Omega := \mathcal{Q}^{n-s}_{\varepsilon_1^{2^s}} \times \mathit{Res}(\mathcal{Q}^{n-s}, \delta')$ (with the latter defined as in Section 2.6.1), so that, given $R \sim \mathit{Res}(I, \delta')$, we have that $(I(G_1), R) \sim \Omega$.

Let $\mathcal{E}_1$ be the event that there exists a tree $T \subseteq I(G_1) - R$ such that the following hold:

(TR1) $\Delta(T) < D$, and

(TR2) for all $x \in V(I)$, we have that $|N_I(x) \cap V(T)| \geq 4(n-s)/5$.

It follows from Theorem 2.6.1, with $n-s$, $\varepsilon_1^{2^s}$, $\delta', \varnothing$ and $1/5$ playing the roles of n, ε, $\delta, \mathcal{S}$ and ε', respectively, that $\mathbb{P}_\Omega[\mathcal{E}_1] = 1 - o(1)$.

Step 2: Identifying scant molecules. For each $v \in V(I)$, let $\mathcal{M}_v$ denote the vertex molecule $\mathcal{M}_v := \{av : a \in \{0,1\}^s\}$. We say a vertex molecule $\mathcal{M}_v$ is *scant* if there exist some layer L and some vertex $x \in V(\mathcal{M}_v \cap L)$ such that $d_H(x, R_L) < \alpha\delta' n/10$, where R_L is the clone of R in L. Let $\mathcal{E}_2$ be the event that there exists some $x \in V(I)$ such that there are more than S' vertices $v \in B_I^{10\ell}(x)$ satisfying that $\mathcal{M}_v$ is scant. It follows from Lemma 2.7.5 with S' and δ' playing the roles of C and δ that $\mathbb{P}_\Omega[\mathcal{E}_2] < e^{-n}$. Let $\mathcal{E}_1^* := \mathcal{E}_1 \wedge \overline{\mathcal{E}_2}$. Therefore, $\mathbb{P}_\Omega[\mathcal{E}_1^*] = 1 - o(1)$.

Condition on $\mathcal{E}_1^*$ holding. Then, G_1 satisfies the following: there exist a set $R \subseteq V(I)$ and a tree $T \subseteq I(G_1) - R$ such that the following hold:

(T1) $\Delta(T) < D$;

(T2) for all $x \in V(I)$, we have that $|N_I(x) \cap V(T)| \geq 4(n-s)/5$, and

(T3) for every $x \in V(I)$, $B_I^{10\ell}(x)$ contains at most S' vertices v such that $\mathcal{M}_v$ is scant.

Recall this implies clones of T and R satisfying (T1)–(T3) exist simultaneously in each layer of G_1.

Step 3: Finding robust matchings for each slice. Recall from Section 2.2.5 that we will absorb vertices in pairs, where each pair consists of two clones x', x'' of the same vertex $x \in V(I)$. In this step, for each $x \in V(I)$ and for each set of clones of x that may need to be absorbed, we find a pairing of these clones so that we can later build suitable absorbing ℓ-cube pairs for each such pair of clones. We will find this pairing separately for each slice of the vertex molecule $\mathcal{M}_x$. Considering each slice separately has the advantage that the chosen pairs are 'localised'. This will be convenient later when linking up the paths used to absorb these vertices. Accordingly, we now partition the set of layers into sets of consecutive layers as follows. Let

$$q := 2^{10Dk^*} \qquad \text{and let} \qquad t := 2^s/q. \tag{2.7.1}$$

For each $j \in [t]$, let $S_j := \bigcup_{i=(j-1)q+1}^{jq} L_i$. Given any molecule $\mathcal{M}$, we consider the slices $\mathcal{S}_j(\mathcal{M}) := S_j \cap \mathcal{M}$. We denote by $\mathcal{S}(\mathcal{M})$ the collection of all these slices of $\mathcal{M}$.

Let $V_{\text{sc}} \subseteq V(I)$ be the set of all vertices $x \in V(I)$ such that $\mathcal{M}_x$ is scant. Recall $G_2 \sim \mathcal{Q}^n_{\varepsilon_2}$. For each $v \in V(I) \setminus V_{\text{sc}}$ and each $\mathcal{S} \in \mathcal{S}(\mathcal{M}_v)$, we define the following auxiliary bipartite graphs. Let $H(\mathcal{S}) := (V(\mathcal{S}), N_I(v), E_H)$, where E_H is defined as follows. Consider $v' \in V(\mathcal{S})$ and let $L^{v'}$ be the layer which contains v'. Let $w \in N_I(v)$, and let $w_{L^{v'}}$ be the clone of w in $L^{v'}$. Then, $\{v', w\} \in E_H$ if and only if $w \in R$ and $\{v', w_{L^{v'}}\} \in E(H)$. Note that $d_{H(\mathcal{S})}(v') \geq \alpha\delta' n/10$ for all $v' \in V(\mathcal{S})$ since $\mathcal{S}$ is a slice of a vertex molecule which is not scant. Similarly, we define $G_2(\mathcal{S}) := (V(\mathcal{S}), N_I(v), E_{G_2})$, where $\{v', w\} \in E_{G_2}$ if and only if $\{v', w_{L^{v'}}\} \in E(G_2)$.

Note that the partition of $V(\mathcal{S})$ into vertices of even and odd parity is a balanced bipartition. Define the graph $\Gamma^{\beta}_{H(\mathcal{S}),G_2(\mathcal{S})}(V(\mathcal{S}))$ as in Section 2.5.1. Note that, by definition, we have that $V(\Gamma^{\beta}_{H(\mathcal{S}),G_2(\mathcal{S})}(V(\mathcal{S}))) = V(\mathcal{S})$. Furthermore, by definition,

(RM) given any $w_1, w_2 \in V(\mathcal{S})$, we have that $\{w_1, w_2\} \in E(\Gamma^{\beta}_{H(\mathcal{S}),G_2(\mathcal{S})}(V(\mathcal{S})))$ if and only if $|N_{H(\mathcal{S})}(w_1) \cap N_{G_2(\mathcal{S})}(w_2)| \geq \beta(n-s)$ or $|N_{G_2(\mathcal{S})}(w_1) \cap N_{H(\mathcal{S})}(w_2)| \geq \beta(n-s)$.

By applying Lemma 2.5.2 with $d = 24D$, $r = 0$, $\alpha = \alpha\delta'/10$, $\varepsilon = \varepsilon_2$, $n = n-s$, $k = q = 2^{10Dk^*}$, $\beta = \beta$ and $G = H(\mathcal{S})$, we obtain that, with probability at least $1 - 2^{-10(n-s)} \geq 1 - 2^{-8n}$, the graph $\Gamma^{\beta}_{H(\mathcal{S}),G_2(\mathcal{S})}(V(\mathcal{S}))$ is $24D$-robust-parity-matchable with respect to the partition of $V(\mathcal{S})$ into vertices of even and odd parity.

We would like to proceed as above for slices in scant molecules; however, recall that scant molecules contain vertices with few or no neighbours in the reservoir, and therefore we must adapt our approach. For each $v \in V_{\mathrm{sc}}$ and each $\mathcal{S} \in \mathcal{S}(\mathcal{M}_v)$, we define an auxiliary bipartite graph $H(\mathcal{S})$ and $G_2(\mathcal{S})$ as above, except that we omit the condition that $w \in R$ for the existence of an edge in $H(\mathcal{S})$. By applying Lemma 2.5.2 again, we obtain that, with probability at least $1 - 2^{-8n}$, the graph $\Gamma^{\beta}_{H(\mathcal{S}),G_2(\mathcal{S})}(V(\mathcal{S}))$ is $24D$-robust-parity-matchable with respect to the partition of $V(\mathcal{S})$ into vertices of even and odd parity.

By a union bound over all $v \in V(I)$ and all slices $\mathcal{S} \in \mathcal{S}(\mathcal{M}_v)$, we have that a.a.s. the graph $\Gamma^{\beta}_{H(\mathcal{S}),G_2(\mathcal{S})}(V(\mathcal{S}))$ is $24D$-robust-parity-matchable (with respect to the partition of $V(\mathcal{S})$ into vertices of even and odd parity) for every slice $\mathcal{S}$, where $H(\mathcal{S})$ is as defined above in each case. We condition on this event holding and call it $\mathcal{E}_2^*$. Thus, for each slice $\mathcal{S}$ and each set $S \subseteq V(\mathcal{S})$ with $|S| \leq 24D$ which contains as many odd vertices as even vertices, there exists a perfect matching $\mathfrak{M}(\mathcal{S}, S)$ in the bipartite graph with parts consisting of the even and odd vertices of $V(\mathcal{S}) \setminus S$, respectively, and edges given by $\Gamma^{\beta}_{H(\mathcal{S}),G_2(\mathcal{S})}(V(\mathcal{S}))$. For each slice $\mathcal{S}$, we denote by $\mathfrak{M}(\mathcal{S})$ the set of edges contained in the union (over all S) of the matchings $\mathfrak{M}(\mathcal{S}, S)$ (without multiplicity). Furthermore, for each $e = \{w_{\mathrm{e}}, w_{\mathrm{o}}\} \in \mathfrak{M}(\mathcal{S})$, we let $N(e) \coloneqq (N_{H(\mathcal{S})}(w_{\mathrm{e}}) \cap N_{G_2(\mathcal{S})}(w_{\mathrm{o}})) \cup (N_{G_2(\mathcal{S})}(w_{\mathrm{e}}) \cap N_{H(\mathcal{S})}(w_{\mathrm{o}}))$.

By (RM), we have $|N(e)| \geq \beta(n-s) \geq \beta n/2$. For each $v \in V(I)$, let $\mathfrak{M}(v) := \bigcup_{\mathcal{S}\in\mathcal{S}(\mathcal{M}_v)} \mathfrak{M}(\mathcal{S})$. Let $K := \max_{v\in V(I)} |\mathfrak{M}(v)|$. In particular, we have that $K \leq \binom{2^s}{2}$.

Step 4: Obtaining an appropriate cube factor via the nibble. For each $x \in V(I)$, consider the multiset $\mathfrak{A}(x) := \{N(e) : e \in \mathfrak{M}(x)\}$. If $|\mathfrak{A}(x)| < K$, we artificially increase its size to K by repeating any of its elements. Label the sets in $\mathfrak{A}(x)$ arbitrarily as $\mathfrak{A}(x) = \{A_1(x), \dots, A_K(x)\}$. Thus, if $x \in V(I) \setminus V_{\mathrm{sc}}$, then $A_i(x) \subseteq R$ for all $i \in [K]$.

Let $\mathcal{C}$ be any collection of subgraphs C of I such that $C \cong \mathcal{Q}^\ell$ for all $C \in \mathcal{C}$. For any vertex $x \in V(I)$ and any set $Y \subseteq N_I(x)$, let $\mathcal{C}_x(Y) \subseteq \mathcal{C}$ be the set of all $C \in \mathcal{C}$ such that $x \notin V(C)$ and $Y \cap V(C) \neq \varnothing$, and let $\mathcal{C}_x := \mathcal{C}_x(N_I(x))$.

Recall $G_3 \sim \mathcal{Q}^n_{\varepsilon_3}$ and $I(G_3) \sim \mathcal{Q}^{n-s}_{\varepsilon_3^{2^s}}$. We now apply Theorem 2.5.7 to the graph $I(G_3)$, with $\varepsilon_3^{2^s}, \alpha', \delta/2, \beta/2, K$ and ℓ playing the roles of $\varepsilon, \alpha, \delta, \beta, K$ and ℓ, respectively, and using the sets $A_i(x)$ given above, for each $x \in V(I)$ and $i \in [K]$. Thus, a.a.s. we obtain a collection $\mathcal{C}$ of vertex-disjoint copies of $\mathcal{Q}^\ell$ in $I(G_3)$, such that the following properties hold for every $x \in V(I)$:

(N1) $|\mathcal{C}_x| \geq (1-\delta)n$.

(N2) For every direction $\hat{e} \in \mathcal{D}(I)$ we have that $|\Sigma(\mathcal{C}_x, \{\hat{e}\}, 1)| = o(n^{1/2})$.

(N3) For every $i \in [K]$ and every $S \subseteq \mathcal{D}(I)$ with $\alpha'(n-s)/2 \leq |S| \leq \alpha'(n-s)$ we have

$$|\Sigma(\mathcal{C}_x(A_i(x)), S, \ell^{1/2})| \geq |A_i(x)|/3000 \geq \beta n/6000.$$

Condition on the above event holding and call it $\mathcal{E}_3^*$.

Step 5: Absorption cubes. For each $x \in V(I)$ and $i \in [K]$, we define an auxiliary digraph $\mathfrak{D} = \mathfrak{D}(A_i(x))$ on vertex set $A_i(x) - \{x\}$ (seen as a set of directions of $\mathcal{D}(I)$) by adding a directed edge from $\hat{e}$ to $\hat{e}'$ if there is a cube $C^r \in \mathcal{C}_x(A_i(x))$ such that $x + \hat{e} \in V(C^r)$ and $\hat{e}' \in \mathcal{D}(C^r)$. In this way, an edge from $\hat{e}$ to $\hat{e}'$ in $\mathcal{D}$ indicates that the cube C^r could be used as a right absorber cube for x, if combined with a vertex-disjoint left absorber cube with tip $x + \hat{e}'$. Observe that, for all $\hat{e} \in A_i(x) - \{x\}$,

$$d^+_{\mathfrak{D}}(\hat{e}) \in [\ell]_0. \tag{2.7.2}$$

Furthermore, it follows by (N3) that any set $S \subseteq V(\mathfrak{D})$ with $|S| = \alpha' n/2$ satisfies

$$e_{\mathfrak{D}}(V(\mathfrak{D}), S) \geq \ell^{1/2}\beta n/6000 > \ell^{1/2}\beta^2 n. \tag{2.7.3}$$

Recall that $A_i(x) = N(\{x_1, x_2\})$ for some $\{x_1, x_2\} \in \mathfrak{M}(\mathcal{S})$, where $\mathcal{S} \in \mathcal{S}(\mathcal{M}_x)$ is some slice of $\mathcal{M}_x$. Note that $x_1, x_2 \in \mathcal{M}_x$, and let L^j be the layer containing x_j for each $j \in [2]$. We say that x_1 and x_2 are the vertices (or clones of x) which *correspond* to the pair (x, i). Let $(\hat{e}, \hat{e}') \in E(\mathfrak{D})$ and, for each $j \in [2]$, let e_j be the clone of $\{x + \hat{e}', x + \hat{e}' + \hat{e}\}$ in L^j. It follows that there is a cube $C^r \in \mathcal{C}_x(A_i(x))$ such that e_j connects the clone C_j of C^r to the clone of $x + \hat{e}'$ in L^j.

Recall $G_4 \sim \mathcal{Q}^n_{\varepsilon_4}$. Let $\mathfrak{D}' \subseteq \mathfrak{D}$ be the subdigraph which retains each edge $(\hat{e}, \hat{e}') \in E(\mathfrak{D})$ if and only if the edges e_1, e_2 described above are both present in G_4. Note that each edge of $\mathfrak{D}$ is therefore retained independently of every other edge with probability ε_4^2. By Lemma 2.4.2, (2.7.2) and (2.7.3), it follows that $\mathfrak{D}'$ satisfies the following with probability at least $1 - e^{-10n}$:

(DG1) for every $A \subseteq V(\mathfrak{D})$ with $|A| = \alpha' n/2$ we have $\sum_{v\in A} d^-_{\mathfrak{D}'}(v) \geq \varepsilon_4^3\beta^2\ell^{1/2}n$, and

(DG2) for every $B \subseteq V(\mathfrak{D})$ we have that $\sum_{v\in B} d^+_{\mathfrak{D}'}(v) \leq \ell|B|$.

Recall that $\mathfrak{D} = \mathfrak{D}(A_i(x))$. By a union bound, (DG1) and (DG2) hold a.a.s. for all $x \in V(I)$ and $i \in [K]$. We condition on this event and call it $\mathcal{E}_4^*$.

For each $x \in V(I)$ and $i \in [K]$, recall that (RM) and the definition of $A_i(x)$ imply that $|A_i(x)| \geq \beta(n - s)$. Thus, it follows by Lemma 2.5.3 with $|A_i(x)|$, $2\alpha'/\beta$, $\varepsilon_4^3\beta^3\ell^{1/2}/(2\alpha')$ and ℓ playing the roles of n, α, c and C, respectively, that there exists a matching $M''(A_i(x))$ of size at least $\frac{\varepsilon_4^3\beta^2}{2\ell^{1/2}}|A_i(x)| \geq \varepsilon_4^3\beta^3 n/(3\ell^{1/2})$ in $\mathfrak{D}'(A_i(x))$.

Next, for each $x \in V(I)$ and $i \in [K]$, we remove from $M''(A_i(x))$ all edges $(\hat{e}, \hat{e}') \in M''(A_i(x))$ such that $x + \hat{e}'$ does not lie in any cube of $\mathcal{C}_x(A_i(x))$. We denote the resulting matching by $M'(A_i(x))$. Note that, by (N1), we have

$$|M'(A_i(x))| \geq \varepsilon_4^3 \beta^3 n/(3\ell^{1/2}) - \delta n \geq n/\ell. \tag{2.7.4}$$

Consider $A_i(x)$, for some $x \in V(I)$ and $i \in [K]$, and let x_1, x_2 be the clones of x which correspond to (x, i). As before, for each $j \in [2]$, let L^j be the layer containing x_j. Recall Definition 2.7.2 and note that, by construction, we have the following.

(AB1) For each edge $(\hat{e}, \hat{e}') \in M'(A_i(x))$, there is an absorbing ℓ-cube pair (C^l, C^r) for x in I such that, for each $j \in [2]$, the clone (C^l_j, C^r_j) of (C^l, C^r) in L^j is an absorbing ℓ-cube pair for x_j in $H \cup G_2 \cup G_3 \cup G_4$. In particular, the edge joining the left absorber tip to the third absorber vertex lies in G_4. Moreover, $C^l, C^r \in \mathcal{C}_x(A_i(x)) \subseteq \mathcal{C}$ and (C^l, C^r) has left and right absorber tips $x + \hat{e}'$ and $x + \hat{e}$, respectively. Furthermore, for each $x \in V(I) \setminus V_{\mathrm{sc}}$, these tips lie in R. We refer to (C^l_1, C^r_1) and (C^l_2, C^r_2) as the absorbing ℓ-cube pairs for x_1 and x_2 *associated* with $(\hat{e}, \hat{e}')$.

Thus, the graph $H \cup G_2 \cup G_3 \cup G_4$, contains at least n/ℓ absorbing ℓ-cube pairs for each of the clones x_1 and x_2 of x associated with edges in $M'(A_i(x)) \subseteq \mathfrak{D}(A_i(x))$. Moreover, since $M'(A_i(x))$ is a matching, for each $j \in [2]$ these absorbing ℓ-cube pairs for x_j are pairwise vertex-disjoint apart from x_j.

For ease of notation, we will often consider the absorbing ℓ-cube pair (C^l, C^r) for x in I which (C^l_1, C^r_1) and (C^l_2, C^r_2) are clones of, and use it as a placeholder for both of its clones. By slightly abusing notation, we will refer to (C^l, C^r) as the *absorbing ℓ-cube pair associated* with $(\hat{e}, \hat{e}')$. Note, however, that (C^l, C^r) is not necessarily an absorbing ℓ-cube pair for x in $I(H \cup G_2 \cup G_3 \cup G_4)$.

Step 6: Removing bondless molecules. Recall $G_5 \sim \mathcal{Q}^n_{\varepsilon_5}$. In this step, we consider the edges between the different layers.

For each $C \in \mathcal{C}$, let $\mathcal{M}_C$ denote the cube molecule consisting of the clones of C. Let $\mathcal{C}' \subseteq \mathcal{C}$ be the set of cubes $C \in \mathcal{C}$ for which $\mathcal{M}_C$ is bonded in G_5. By an application of Lemma 2.7.3, for each $C \in \mathcal{C}$ we have that

$$\mathbb{P}[C \notin \mathcal{C}'] = \mathbb{P}[\mathcal{M}_C \text{ is bondless in } G_5] \leq 2^{s+1-\varepsilon_5 2^\ell/4} \leq 2^{-\varepsilon 2^\ell/30}.$$

For each $x \in V(I)$, let $A_0(x) \coloneqq N_I(x)$. For each $i \in [K]_0$, let $\mathcal{E}(x,i)$ be the event that $|\mathcal{C}_x(A_i(x)) \setminus \mathcal{C}'| > n/\ell^4$. Since the cubes $C \in \mathcal{C}$ are vertex-disjoint, the events that the molecules $\mathcal{M}_C$ are bondless in G_5 are independent. Therefore, we have that

$$\mathbb{P}[\mathcal{E}(x,i)] \leq \binom{n}{n/\ell^4}(2^{-\varepsilon 2^\ell/30})^{n/\ell^4} \leq 2^{-10n}.$$

Let $\mathcal{E}_4 \coloneqq \bigvee_{x\in V(I)} \bigvee_{i\in[K]_0} \mathcal{E}(x,i)$. By a union bound over all $x \in V(I)$ and $i \in [K]_0$, it follows that

$$\mathbb{P}[\mathcal{E}_4] \leq 2^{-8n}. \tag{2.7.5}$$

Let $\mathcal{C}_{\text{bs}} \subseteq \mathcal{C}$ be the set of all $C \in \mathcal{C}$ such that $\mathcal{M}_C$ is bondlessly surrounded in G_5 (with respect to $\{\mathcal{M}_{C'} : C' \in \mathcal{C}\}$). For each $x \in V(I)$, let $\mathcal{E}(x)$ be the event that there are more than $n^{1/3}$ cubes $C \in \mathcal{C}_{\text{bs}}$ which intersect $B_I^{\ell^2}(x)$. Let $\mathcal{E}_5 \coloneqq \bigvee_{x\in V(I)} \mathcal{E}(x)$. By (N2), we may apply Lemma 2.7.4 with ε_5 playing the role of ε to conclude that

$$\mathbb{P}[\mathcal{E}_5] \leq 2^{-n^{9/8}}. \tag{2.7.6}$$

Now let $\mathcal{E}_5^* \coloneqq \overline{\mathcal{E}_4} \wedge \overline{\mathcal{E}_5}$. It follows from (2.7.5) and (2.7.6) that $\mathcal{E}_5^*$ occurs a.a.s. Condition on this event.

Let $\mathcal{C}'' \coloneqq \mathcal{C}' \setminus \mathcal{C}_{\text{bs}}$. For each $x \in V(I)$ and each $i \in [K]$, let

(AB2) $M(A_i(x)) \subseteq M'(A_i(x))$ consist of all edges $(\hat{e}, \hat{e}') \in M'(A_i(x))$ whose associated absorbing ℓ-cube pair (C^l, C^r) satisfies that $C^r, C^l \in \mathcal{C}''$.

By combining (2.7.4) with the further conditioning, it follows that, for each $x \in V(I)$ and each $i \in [K]$,

$$|M(A_i(x))| \geq n/\ell - n/\ell^4 - n^{1/3} \geq n/\ell^2. \tag{2.7.7}$$

Consider any $x \in V(I)$ and $i \in [K]$, and let x_1, x_2 be the two clones of x corresponding to (x, i). Then, at this point, for each $j \in [2]$, $H \cup G_2 \cup G_3 \cup G_4$ contains at least n/ℓ^2 vertex-disjoint (apart from x_j) absorbing ℓ-cube pairs for x_j such that each of these absorbing ℓ-cube pairs (C^l, C^r) is associated with an edge of $M(A_i(x))$, and for each $C \in \{C^l, C^r\}$ the corresponding cube molecule $\mathcal{M}_C$ is bonded in G_5 and (within the collection $\{\mathcal{M}_{C'} : C' \in \mathcal{C}\}$ of all cube molecules) $\mathcal{M}_C$ is not bondlessly surrounded in G_5.

Step 7: Extending the tree T. For each $x \in V(I)$, let $Z(x) \coloneqq N_I(x) \cap V(T) \cap \left(\bigcup_{C \in \mathcal{C}''} V(C)\right)$. It follows by (T2), (N1) and our conditioning on the event $\mathcal{E}_5^*$ that, for each $x \in V(I)$, we have that

$$|Z(x)| \geq 4(n-s)/5 - \delta n - n/\ell^4 - n^{1/3} \geq 3n/4.$$

Recall $G_6 \sim \mathcal{Q}^n_{\varepsilon_6}$. We apply Theorem 2.6.19 with $\varepsilon_6^{2^s}$, 2, T, $R, \varnothing$ and the sets $Z(x)$ playing the roles of ε, ℓ, T', R, W and $Z(x)$, respectively. Combining this with (T1), we conclude that a.a.s. there exists a tree T' such that $T \subseteq T' \subseteq I(G_6) \cup T$ and the following hold:

(ET1) $\Delta(T') < D + 1$;

(ET2) for all $x \in V(I)$, we have that $|B_I^2(x) \setminus V(T')| \leq n^{3/4}$;

(ET3) for each $x \in V(T') \cap R$, we have that $d_{T'}(x) = 1$ and the unique neighbour x' of x in T' is such that $x' \in Z(x)$.

We condition on the above event holding and call it $\mathcal{E}_6^*$.

At this point, for each $x \in V(I)$ and each $i \in [K]$, we redefine the set $M(A_i(x))$.

(AB3) Let $M(A_i(x))$ retain only those edges whose associated absorbing ℓ-cube pair (C^l, C^r) satisfies that both C^l and C^r intersect T' in at least 2 vertices.

It follows from (2.7.7) and (ET2) that

$$|M(A_i(x))| \geq n/\ell^2 - n^{3/4} > 4n/\ell^3. \tag{2.7.8}$$

Step 8: Fixing a collection of absorbing ℓ-cube pairs for the vertices in scant molecules. Recall $G_7 \sim \mathcal{Q}^n_{\varepsilon_7}$. Consider any $x \in V_{\rm sc}$ and $j \in [K]$. Recall from Step 3 that the tips of the cubes of the absorbing ℓ-cube pair associated with a given edge in $M(A_j(x))$ may not lie in the reservoir R. Roughly speaking, we will alter T' so that the tips are relocated from the tree T' to the reservoir R.

We start by redefining the matchings $M(A_j(x))$ as follows: for each $x \in V_{\rm sc}$ and each $j \in [K]$, remove from $M(A_j(x))$ all edges $(\hat{e}, \hat{e}')$ such that $N_{T'}(x) \cap \{x+\hat{e}, x+\hat{e}'\} \neq \varnothing$. It follows from (2.7.8) and (ET1) that, for all $x \in V(I)$ and $j \in [K]$,

$$|M(A_j(x))| \geq 4n/\ell^3 - D > 2n/\ell^3. \tag{2.7.9}$$

For each $x \in V_{\rm sc}$, each $j \in [K]$ and each matching $M' \subseteq M(A_j(x))$ with $|M'| \geq n/\ell^3$, let $\mathcal{E}'(x, j, M')$ be the following event:

> for every set $B \subseteq V(I)$ with $|B| < 2^{\ell+s+3}\Psi K S'$, there exists an edge $\vec{e} \in M'$, whose associated absorbing ℓ-cube pair (C^l, C^r) has tips x^l and x^r, for which there exists a subgraph $P(\vec{e}, B) \subseteq I(G_7) - \{x^l, x^r\}$ such that $|V(P(\vec{e}, B))| < 21D/2$, $V(P(\vec{e}, B)) \cap B = \varnothing$, and both $N_{T'}(x^l)$ and $N_{T'}(x^r)$ are connected in $P(\vec{e}, B)$.

For a graph $P(\vec{e}, B)$ as above, we will refer to x^l and x^r as the tips *associated* with $P(\vec{e}, B)$, and refer to (C^l, C^r) as the absorbing ℓ-cube pair *associated* with $P(\vec{e}, B)$. (Recall that, if $\vec{e} = (\hat{e}, \hat{e}')$, then $x^l = x + \hat{e}$ and $x^r = x + \hat{e}'$.)

By invoking Lemma 2.6.20 with $n - s$, $\varepsilon_7^{2^s}$, $1/\ell^3$, $2^{\ell+s+3}\Psi K S'$, $2D + 2$ and the sets $\{(x+\hat{e}, x+\hat{e}') : (\hat{e}, \hat{e}') \in M'\}$ and $(N_{T'}(x+\hat{e}) \cup N_{T'}(x+\hat{e}'))_{(\hat{e},\hat{e}') \in M'}$ playing the roles of n, ε, c, f, D, $C(x)$ and $(B(y,z))_{(y,z)\in C(x)}$, respectively, we have that $\mathcal{E}'(x, j, M')$ holds with probability at least $1 - 2^{-5(n-s)}$. Let $\mathcal{E}^*_7 \coloneqq \bigwedge_{x \in V_{\rm sc}} \bigwedge_{j \in [K]} \bigwedge_{M' \subseteq M(A_j(x)) : |M'| \geq n/\ell^3} \mathcal{E}'(x, j, M')$.

By a union bound over all $x \in V_{\rm sc}$, $j \in [K]$ and $M' \subseteq M(A_j(x))$ such that $|M'| \geq n/\ell^3$, it follows that $\mathbb{P}[\mathcal{E}_7^*] \geq 1 - 2^{-2n}$.

Condition on the event that $\mathcal{E}_7^*$ holds. It follows that, for each $x \in V_{\rm sc}$, $j \in [K]$, $M' \subseteq M(A_j(x))$ with $|M'| \geq n/\ell^3$ and any $B \subseteq V(I)$ with $|B| < 2^{\ell+s+3}\Psi K S'$, there exists a subgraph $P(x, j, M', B) \subseteq I(G_7)$ with $|V(P(x, j, M', B))| < 21D/2$ which avoids $B \cup \{x^l, x^r\}$, where x^l and x^r are the tips associated with $P(x, j, M', B)$, and such that both $N_{T'}(x^l)$ and $N_{T'}(x^r)$ are connected in $P(x, j, M', B)$. Moreover, by choosing $P(x, j, M', B)$ minimal, we may assume that it consists of at most two components, and each such component contains either $N_{T'}(x^l)$ or $N_{T'}(x^r)$.

Let $\iota := |V_{\rm sc}|$ and let $x_1, \ldots, x_\iota$ be an ordering of $V_{\rm sc}$. For each $i \in [\iota]$, $j \in [K]$ and $k \in [2^{s+1}\Psi]$, by ranging over i first, then j, and then k, we will iteratively fix a graph $P(x_i, j, k, M'_{i,j,k}, B_{i,j,k})$ as above. In particular, this graph will have an absorbing ℓ-cube pair with tips $x^l_{i,j,k}$ and $x^r_{i,j,k}$ associated with it. After the graph $P(x_i, j, k, M'_{i,j,k}, B_{i,j,k})$ is fixed, so are these tips. Let $\mathcal{J}_{i,j,k} := ([i-1] \times [K] \times [2^{s+1}\Psi]) \cup \{(i, j', k') : (j', k') \in [j-1] \times [2^{s+1}\Psi]\} \cup \{(i, j, k'') : k'' \in [k-1]\}$ and suppose that we have already fixed $P(x_{i'}, j', k', M'_{i',j',k'}, B_{i',j',k'})$ for all $(i', j', k') \in \mathcal{J}_{i,j,k}$ such that these $P(x_{i'}, j', k', M'_{i',j',k'}, B_{i',j',k'})$ are vertex-disjoint from each other and from the set $\{x^l_{i',j',k'}, x^r_{i',j',k'} : (i', j', k') \in \mathcal{J}_{i,j,k}\}$ of tips associated with all these $P(x_{i'}, j', k', M'_{i',j',k'}, B_{i',j',k'})$. In order to fix $P(x_i, j, k, M'_{i,j,k}, B_{i,j,k})$, we first define the sets $B_{i,j,k}$ and $M'_{i,j,k}$. Let $M'_{i,j,k}$ be obtained from $M(A_j(x_i))$ as follows. Remove all edges whose associated absorbing ℓ-cube pair (C^l, C^r) satisfies $(V(C^l) \cup V(C^r)) \cap \{x^l_{i',j',k'}, x^r_{i',j',k'} : (i', j', k') \in \mathcal{J}_{i,j,k}\} \neq \varnothing$. Remove all edges $(\hat{e}, \hat{e}') \in M(A_j(x_i))$ such that $\{x_i + \hat{e}, x_i + \hat{e}'\} \cap \bigcup_{(i',j',k') \in \mathcal{J}_{i,j,k}} V(P(x_{i'}, j', k', M'_{i',j',k'}, B_{i',j',k'})) \neq \varnothing$ too. Note that, by (2.7.9) and (T3), it follows that $|M'_{i,j,k}| \geq n/\ell^3$. Let $B_{i,j,k}$ be the set of vertices $y \in B_I^{\ell/2}(x_i)$ such that at least one of the following holds:

(P1) there exists $(i', j', k') \in \mathcal{J}_{i,j,k}$ such that $y \in V(P(x_{i'}, j', k', M'_{i',j',k'}, B_{i',j',k'}))$;

(P2) there exists $(i', j', k') \in \mathcal{J}_{i,j,k}$ such that y lies in the absorbing ℓ-cube pair associated with $P(x_{i'}, j', k', M'_{i',j',k'}, B_{i',j',k'})$.

Note that $|B_{i,j,k}| < 2^{s+\ell+3}\Psi KS'$ by (T3). We then fix $P(x_i, j, k, M'_{i,j,k}, B_{i,j,k})$ to be the graph guaranteed by our conditioning on $\mathcal{E}_7^*$. Observe that, by the choice of $B_{i,j,k}$, we have that $P(x_i, j, k, M'_{i,j,k}, B_{i,j,k})$ is vertex-disjoint from $\bigcup_{(i',j',k')\in\mathcal{J}_{i,j,k}} P(x_{i'}, j', k', M'_{i',j',k'}, B_{i',j',k'})$. We denote by $(C^l(x_i, j, k), C^r(x_i, j, k))$ the absorbing ℓ-cube pair for x_i associated with $P(x_i, j, k, M'_{i,j,k}, B_{i,j,k})$. By the choice of $M'_{i,j,k}$, we have that

(CD) for all $(i', j', k') \in \mathcal{J}_{i,j,k}$, $C^l(x_i, j, k)$ and $C^r(x_i, j, k)$ are both vertex-disjoint from $C^l(x_{i'}, j', k')$ and $C^r(x_{i'}, j', k')$.

Let $\mathcal{C}_1^{\mathrm{sc}} := \{(C^l(x_i, j, k), C^r(x_i, j, k)) : (i, j, k) \in [\iota] \times [K] \times [2^{s+1}\Psi]\}$. Let $P' := \{x^l_{i,j,k}, x^r_{i,j,k} : (i, j, k) \in [\iota] \times [K] \times [2^{s+1}\Psi]\}$ and $P := \bigcup_{i\in[\iota], j\in[K], k\in[2^{s+1}\Psi]} P(x_i, j, k, M'_{i,j,k}, B_{i,j,k})$. Recall that $P(x_i, j, k, M'_{i,j,k}, B_{i,j,k})$ avoids the tips $x^l_{i,j,k}$ and $x^r_{i,j,k}$ associated with it. It follows from this, (P2), and the definition of $M'_{i,j,k}$ that $P' \cap V(P) = \varnothing$. Let $T''' := T'[V(T') \setminus P'] \cup P$. Note that T''' is connected by the definition of $\mathcal{E}'(x, j, M')$. Let T'' be a spanning tree of T'''. By (ET1) and the fact that the graphs $P(x_i, j, k, M'_{i,j,k}, B_{i,j,k})$ are vertex-disjoint and satisfy $|V(P(x_i, j, k, M'_{i,j,k}, B_{i,j,k}))| < 21D/2$, it follows that

$$\Delta(T'') \le 12D. \tag{2.7.10}$$

Define the (new) reservoir $R' := (R \cup P') \setminus V(P)$.

At this point, for each $x \in V(I) \setminus V_{\mathrm{sc}}$ and each $i \in [K]$, we redefine the set $M(A_i(x))$ as follows.

(AB4) Let $M(A_i(x))$ retain only those edges whose associated absorbing ℓ-cube pair (C^l, C^r) satisfies that both C^l and C^r are vertex-disjoint from both cubes of all absorbing ℓ-cube pairs of $\mathcal{C}_1^{\mathrm{sc}}$ and both tips x^l and x^r satisfy that $x^l, x^r \in R \setminus V(P) \subseteq R'$.

Note that, by (T3), we have $|B_I^{\ell+1}(x) \cap V(P)| \le 21 \cdot 2^s \Psi DKS'$ and $|B_I^{\ell+1}(x) \cap V(\bigcup_{(C^l,C^r)\in\mathcal{C}_1^{\mathrm{sc}}}(C^l \cup C^r))| \le 4 \cdot 2^{\ell+s}\Psi KS'$. Combining this with (2.7.8) and (AB1), it follows that

$$|M(A_i(x))| \ge 4n/\ell^3 - (21D + 4 \cdot 2^\ell)2^s\Psi KS' > n/\ell^3. \tag{2.7.11}$$

Step 9: Fixing a collection of absorbing ℓ-cube pairs for the vertices in non-scant molecules. At this point, we still do not know which vertices will need to be absorbed eventually into an almost spanning cycle, but we can already determine the vertices in I whose clones the vertices to be absorbed will be (the reason for this will be apparent later, see Step 13). Recall that $\mathcal{C}'$ and $\mathcal{C}''$ were defined in Step 6. Let $\mathcal{C}''' \coloneqq \{C \in \mathcal{C}' : V(C) \cap V(T'') \neq \varnothing\}$ and let $V_{\mathrm{abs}} \coloneqq V(I) \setminus \bigcup_{C \in \mathcal{C}'''} V(C)$. We will now fix a collection of absorbing ℓ-cube pairs for all vertices in each vertex molecule $\mathcal{M}_x$ with $x \in V_{\mathrm{abs}} \setminus V_{\mathrm{sc}}$.

First, recall from (T3) that, for all $x \in V(I)$, we have that $|B_I^{10\ell}(x) \cap V_{\mathrm{sc}}| \leq S'$. Thus, in constructing T'', we removed at most $2^{s+2}\Psi K S'$ vertices in $B_I^{\ell}(x)$ from T'. Therefore, it follows from (ET2) that, for all $x \in V(I)$, we have

$$|B_I^2(x) \setminus V(T'')| \leq 2n^{3/4}. \tag{2.7.12}$$

For all $x \in \bigcup_{C \in \mathcal{C}''} V(C)$, we claim that

$$|N_I(x) \cap V(T'') \cap \bigcup_{C \in \mathcal{C}'} V(C)| \geq (1 - 2^{1-\ell-5s})n. \tag{2.7.13}$$

To see that this holds, combine (N1), (2.7.12) and the definition of bondlessly surrounded molecules.

Recall also the definition of $\mathfrak{M}(x)$ from Step 3.

Claim 2.7.2. *For each $x \in V_{\mathrm{abs}} \setminus V_{\mathrm{sc}}$ and each $e \in \mathfrak{M}(x)$, there exists a set $\mathcal{C}_1^{\mathrm{abs}}(e)$ of $2^{s+1}\Psi$ absorbing ℓ-cube pairs $(C_k^l(e), C_k^r(e)) \subseteq I$, for $k \in [2^{s+1}\Psi]$, which satisfies the following:*

(i) *for all $x \in V_{\mathrm{abs}} \setminus V_{\mathrm{sc}}$, $e \in \mathfrak{M}(x)$ and $k \in [2^{s+1}\Psi]$, the absorbing ℓ-cube pair $(C_k^l(e), C_k^r(e))$ is associated with some edge in $M(A_j(x))$, for some $j \in [K]$, and*

(ii) *for all $x, x' \in V_{\mathrm{abs}} \setminus V_{\mathrm{sc}}$, all $e \in \mathfrak{M}(x)$ and $e' \in \mathfrak{M}(x')$, and all $k, k' \in [2^{s+1}\Psi]$ with $(x, e, k) \neq (x', e', k')$, the absorbing ℓ-cube pairs $(C_k^l(e), C_k^r(e))$ are vertex-disjoint (except for x in the case when $x = x'$).*

Proof. Let $\mathcal{V} := \bigcup_{x \in V_{\text{abs}} \setminus V_{\text{sc}}} \mathfrak{M}(x)$. Let $K' := |\mathcal{V}|$, and let $f_1, \ldots, f_{K'}$ be an ordering of the edges in $\mathcal{V}$. Given any $i \in [K']$, the edge f_i corresponds to a pair $(x, j(i))$ (in the sense that $A_{j(i)}(x) = N(f_i)$, see Step 4), where $x \in V_{\text{abs}} \setminus V_{\text{sc}}$ and $j(i) \in [K]$. Let $\mathfrak{C}_i$ be the collection of at least n/ℓ^3 absorbing ℓ-cube pairs for x in I guaranteed by (2.7.11). In particular, each of these absorbing ℓ-cube pairs (C^l, C^r) is associated with an edge of $M(A_{j(i)}(x))$ and, by (AB2), satisfies $C^l, C^r \in \mathcal{C}''$.

Let $\mathcal{H}$ be the $2^{s+1}\Psi K'$-edge-coloured auxiliary multigraph with $V(\mathcal{H}) := \mathcal{C}''$, which contains an edge between C and C' of colour $(i, k) \in [K'] \times [2^{s+1}\Psi]$ whenever $(C, C') \in \mathfrak{C}_i$ or $(C', C) \in \mathfrak{C}_i$. In particular, $\mathcal{H}$ contains at least n/ℓ^3 edges of each colour. We now bound $\Delta(\mathcal{H})$. Consider any $C \in V(\mathcal{H})$. Note that, for each edge e of $\mathcal{H}$ incident to C, there exists some $x = x(e) \in V_{\text{abs}} \setminus V_{\text{sc}}$ such that C together with some other cube $C' \in V(\mathcal{H})$ forms an absorbing ℓ-cube pair for x. In particular, x must be adjacent to C in I. Moreover, if e has colour (i, k), then $f_i \in \mathfrak{M}(x)$ (and it has corresponding pair $(x, j(i))$ for some $j(i) \in [K]$). Since $f_i \in \mathfrak{M}(x)$ and $|\mathfrak{M}(x)| \leq \binom{2^s}{2}$, it follows that each vertex y which is adjacent to C in I can play the role of x for at most $2^{s+1}\Psi \cdot 2^{2s}$ edges of $\mathcal{H}$ incident to C. Thus, $d_{\mathcal{H}}(C)$ is at most $2^{s+1}\Psi \cdot 2^{2s}$ times the number of vertices $y \in V_{\text{abs}} \setminus V_{\text{sc}}$ which are adjacent to C in I. Recall that $V_{\text{abs}} = V(I) \setminus \bigcup_{C \in \mathcal{C}'''} V(C)$. Together with (2.7.13), this implies that the number of vertices in V_{abs} which are adjacent to C is at most $|C|n/2^{\ell+5s-1}$. Thus, $d_{\mathcal{H}}(C) \leq 2^{s+1}\Psi 2^{2s}|C|n/2^{\ell+5s-1} \leq n/\ell^4$.

Since each colour class has size at least n/ℓ^3 and $\Delta(\mathcal{H}) \leq n/\ell^4$, by Lemma 2.5.4, $\mathcal{H}$ contains a rainbow matching of size $2^{s+1}\Psi K'$. For each $(i, k) \in [K'] \times [2^{s+1}\Psi]$, let $(C_k^l(f_i), C_k^r(f_i)) \in \mathfrak{C}_i$ be the absorbing ℓ-cube pair of colour (i, k) in this rainbow matching. ◀

Recall that, for any $x \in V(I)$, each index $i \in [K]$ is given by a unique edge $e \in \mathfrak{M}(x)$ via the relation $N(e) = A_i(x)$. For each $x \in V_{\text{abs}} \setminus V_{\text{sc}}$ and each $i \in [K]$, let $\mathcal{C}_1^{\text{abs}}(x, i) := \mathcal{C}_1^{\text{abs}}(e)$, where e is the unique edge given by the relation above, be the set of absorbing ℓ-cube pairs guaranteed by Claim 2.7.2. Similarly, for each $k \in [2^{s+1}\Psi]$, let $(C_k^l(x, i), C_k^r(x, i)) := (C_k^l(e), C_k^r(e))$.

Let $G \coloneqq \bigcup_{i=1}^{7} G_i$. For each $x \in V_{\mathrm{abs}} \setminus V_{\mathrm{sc}}$ and each $i \in [K]$, let $G^*(x,i) \subseteq I$ be the graph consisting of all edges between the left absorber tip and third absorber vertex of every absorbing ℓ-cube pair in $\mathcal{C}_1^{\mathrm{abs}}(x,i)$. Let $G^\bullet \subseteq I$ be the graph consisting of all edges between the left absorber tip and third absorber vertex of every absorbing ℓ-cube pair in $\mathcal{C}_1^{\mathrm{sc}}$. Let $G^* \coloneqq G^\bullet \cup \bigcup_{x \in V_{\mathrm{abs}} \setminus V_{\mathrm{sc}}} \bigcup_{i \in [K]} G^*(x,i) \subseteq I$. Recall that, given any graph $\mathcal{G} \subseteq I$, for each layer L, we denote by $\mathcal{G}_L$ the clone of $\mathcal{G}$ in L. Let $G_4^* \coloneqq G_4 \cap \bigcup_{i=1}^{2^s} G_{L_i}^*$. Furthermore, let $G_5^* \subseteq G_5$ consist of all edges of G_5 which have endpoints in different layers. We let $G' \subseteq G$ be the spanning subgraph with edge set

$$E(G') \coloneqq E(G_4^*) \cup E(G_5^*) \cup \bigcup_{C \in \mathcal{C}'} E(\mathcal{M}_C) \cup \bigcup_{i=1}^{2^s} E(T''_{L_i}).$$

Note that, using (2.7.10), we have that $\Delta(G') \leq \Phi$.

Now, let $F \subseteq \mathcal{Q}^n$ be any graph with $\Delta(F) \leq \Psi$. Recall that we denote by $F_I \subseteq I$ the graph which contains every edge $\{x,y\} \in E(I)$ such that there exists an edge $e = \{x',y'\} \in E(F)$ with $x' \in \mathcal{M}_x$ and $y' \in \mathcal{M}_y$.

Note that $T'' \subseteq I(G')$, $R' \subseteq V(I)$, and $C \subseteq I(G')$ for every $C \in \mathcal{C}'$. Recall the definitions of $\mathcal{C}''$ from Step 6 and $\mathcal{C}'''$ from Step 9. Combining all the previous steps, we claim that the following hold (conditioned on the events $\mathcal{E}_1^*, \ldots, \mathcal{E}_7^*$, which occur a.a.s.).

(C1) $\Delta(T'') \leq 12D$.

(C2) Any vertex $x \in R' \cap V(T'')$ is a leaf of T''. Furthermore, if $x \in R' \cap V(T'')$, then its unique neighbour x' in T'' satisfies that $x' \in Z(x)$ (where $Z(x)$ is as defined in Step 7).

(C3) For all $x \in V(I)$, we have that $|N_I(x) \cap V(T'') \cap \bigcup_{C \in \mathcal{C}''} V(C)| \geq (1 - 2/\ell^4)n$.

(C4) For each $x \in V_{\mathrm{sc}}$ and $i \in [K]$, there is an absorbing ℓ-cube pair $(C^l(x,i), C^r(x,i))$ for x in I, which is associated with some edge $e \in M(A_i(x))$. In particular, $(C^l(x,i), C^r(x,i))$ is as described in (AB1) (recall also (AB2)), that is, there are two absorbing ℓ-cube pairs $(C_1^l(x,i), C_1^r(x,i))$ and $(C_2^l(x,i), C_2^r(x,i))$ in $H \cup G$, associated with $e \in M(A_i(x))$, for the clones x_1 and x_2 of x which correspond to (x,i). Additionally, each of these absorbing ℓ-cube pairs $(C^l(x,i), C^r(x,i))$ satisfies the following:

(C4.1) $(C_1^l(x,i), C_1^r(x,i)) \cup (C_2^l(x,i), C_2^r(x,i)) - V(\mathcal{M}_x) \subseteq G'$;

(C4.2) the tips x^l of $C^l(x,i)$ and x^r of $C^r(x,i)$ lie in $R' \setminus V(T'')$, and $\{x, x^l\}, \{x, x^r\} \notin E(F_I)$; in particular, the tips x_1^l, x_1^r of $(C_1^l(x,i), C_1^r(x,i))$ and x_2^l, x_2^r of $(C_2^l(x,i), C_2^r(x,i))$ satisfy that $\{x_1, x_1^l\}, \{x_1, x_1^r\}, \{x_2, x_2^l\}, \{x_2, x_2^r\} \in E((H \cup G) \setminus F)$;

(C4.3) $C^l(x,i), C^r(x,i) \in \mathcal{C}'' \cap \mathcal{C}'''$, and

(C4.4) for any $x' \in V_{\rm sc}$ and $i' \in [K]$ with $(x',i') \neq (x,i)$ we have that $C^l(x,i)$, $C^r(x,i)$, $C^l(x',i')$ and $C^r(x',i')$ are vertex-disjoint.

Let $\mathcal{C}^{\rm sc}$ denote the collection of these absorbing ℓ-cube pairs.

(C5) For each $x \in V_{\rm abs} \setminus V_{\rm sc}$ and $i \in [K]$, there is an absorbing ℓ-cube pair $(C^l(x,i), C^r(x,i))$ for x in I, which is associated with some edge in $M(A_i(x))$. In particular, $(C^l(x,i), C^r(x,i))$ is as described in (AB1), that is, there are two absorbing ℓ-cube pairs $(C_1^l(x,i), C_1^r(x,i))$ and $(C_2^l(x,i), C_2^r(x,i))$ in $H \cup G$, associated with $e \in M(A_i(x))$, for the clones x_1 and x_2 of x which correspond to (x,i). Moreover, each of these absorbing ℓ-cube pairs $(C^l(x,i), C^r(x,i))$ satisfies the following:

(C5.1) $(C_1^l(x,i), C_1^r(x,i)) \cup (C_2^l(x,i), C_2^r(x,i)) - V(\mathcal{M}_x) \subseteq G'$;

(C5.2) the tips x_i^l of $C^l(x,i)$ and x_i^r of $C^r(x,i)$ lie in R', and $\{x, x_i^l\}, \{x, x_i^r\} \notin E(F_I)$; in particular, the tips x_1^l, x_1^r of $(C_1^l(x,i), C_1^r(x,i))$ and x_2^l, x_2^r of $(C_2^l(x,i), C_2^r(x,i))$ satisfy that $\{x_1, x_1^l\}, \{x_1, x_1^r\}, \{x_2, x_2^l\}, \{x_2, x_2^r\} \in E((H \cup G) \setminus F)$;

(C5.3) $C^l(x,i), C^r(x,i) \in \mathcal{C}'' \cap \mathcal{C}'''$;

(C5.4) for any $x' \in V_{\rm abs} \setminus V_{\rm sc}$ and $i' \in [K]$ with $(x',i') \neq (x,i)$ we have that $C^l(x,i)$, $C^r(x,i)$, $C^l(x',i')$ and $C^r(x',i')$ are vertex-disjoint, and

(C5.5) both $C^l(x,i)$ and $C^r(x,i)$ are vertex-disjoint from all cubes of absorbing ℓ-cube pairs in $\mathcal{C}^{\mathrm{sc}}$.

Let $\mathcal{C}^{\neg\mathrm{sc}}$ denote the collection of these absorbing ℓ-cube pairs.

Indeed, (C1) is given in (2.7.10). (C2) holds by (ET3) and the fact that $P' \cap V(T'') = \varnothing$. (C3) follows by combining (N1), the conditioning on $\mathcal{E}_5^*$, and (2.7.12). (C4) follows from the construction of P and T'' in Step 8. Indeed, for each $x \in V_{\mathrm{sc}}$ and $i \in [K]$, consider the collection of absorbing ℓ-cube pairs $\{(C^l(x,i,k), C^r(x,i,k))\}_{k\in[2^{s+1}\Psi]}$ defined in Step 8. Since $\Delta(F) \leq \Psi$, it follows that $d_{F_I}(x) \leq 2^s\Psi$, and thus there must exist some absorbing ℓ-cube pair in this collection such that the edges joining its tips to x do not belong to F_I. Fix one such absorbing ℓ-cube pair and call it $(C^l(x,i), C^r(x,i))$. Then, (C4.1) holds by the definition of G' combined with (AB1), and (C4.2) holds by the definition of R' and T'' combined with (AB1), while (C4.4) holds by (CD). On the other hand, (C4.3) follows because of the definition of the set $M(A_i(x))$ in (AB2) and (AB3). Finally, consider (C5). For each $x \in V_{\mathrm{abs}} \setminus V_{\mathrm{sc}}$ and $i \in [K]$, consider the collection $\mathcal{C}_1^{\mathrm{abs}}(x,i)$ of $2^{s+1}\Psi$ absorbing ℓ-cube pairs for x in I guaranteed by Claim 2.7.2. For each of these absorbing ℓ-cube pairs we have that (C5.3) holds by (AB2), (AB3) and the fact that, by (AB4), their intersection with T'' contains their intersection with T'. Similarly, (C5.4) holds by Claim 2.7.2, and (C5.5) holds because of (AB4). Finally, note that $\Delta(F_I) \leq 2^s\Psi$. It follows that there exists a choice of $(C^l(x,i), C^r(x,i)) \in \mathcal{C}_1^{\mathrm{abs}}(x,i)$ such that $\{x, x_i^l\}, \{x, x_i^r\} \notin E(F_I)$. Then, (C5.1) and (C5.2) hold by the definition of G', (AB1) and (AB4).

Step 10: Constructing auxiliary trees T^* and τ_0. From this point on, every step will be deterministic. Let T^* be obtained from T'' by removing all leaves of T'' which lie in R'.

We will now construct an auxiliary tree τ_0, which will be used in the construction of an almost spanning cycle. We start by defining an auxiliary multigraph Γ' as follows. First, let $\Gamma_1 \coloneqq T^* \cup \bigcup_{C \in \mathcal{C}'} C$. (Recall that $\mathcal{C}'$ is the collection of all $C \in \mathcal{C}$ for which $\mathcal{M}_C$ is bonded in G_5, see Step 6.) Let Γ_2 be the graph obtained by iteratively removing all leaves from Γ_1 until all vertices have degree at least 2. Observe that, after this is achieved, the resulting graph still contains all cubes $C \in \mathcal{C}'$. Let Γ_3 be obtained from Γ_2 by removing all connected components which consist of a single cube $C \in \mathcal{C}'$. Now, let Γ' be the multigraph obtained by contracting each cube $C \in \mathcal{C}'$ such that $C \subseteq \Gamma_3$ into a single vertex. We refer to the vertices resulting from contracting such cubes as *atomic vertices*, and to the remaining vertices in Γ' as *inner tree vertices*. Given $C \in \mathcal{C}$ and $j \in [2^s]$, we call $\mathcal{A} = \mathcal{M}_C \cap L_j$ an *atom*. We continue to identify each inner tree vertex v with the vertex $v \in V(I)$ from which it originated in Γ_1. Observe that Γ' is connected, and (C1) implies that

$$d_{\Gamma'}(v) \leq 12D \text{ for all inner tree vertices, and } \Delta(\Gamma') \leq 12 \cdot 2^{\ell} D. \tag{2.7.14}$$

Given an atomic vertex $v \in V(\Gamma')$, let $C(v) \in \mathcal{C}$ be the cube which was contracted to v in the construction of Γ', and let $\mathcal{M}(v) \coloneqq \mathcal{M}_{C(v)}$. Furthermore, for each $j \in [2^s]$, let $\mathcal{A}_j(v) \coloneqq \mathcal{M}(v) \cap L_j$. Similarly, for any $v \in V(\Gamma')$ which is an inner tree vertex, we define $\mathcal{M}(v) \coloneqq \mathcal{M}_v$. Observe that every edge $e \in E(\Gamma')$ corresponds to a unique edge $e' \in I(G')$. We say that e *originates* from e'. We denote by $D(e) \in \mathcal{D}(I)$ the direction of e' in I. By abusing notation, we will sometimes also view $D(e)$ as a direction in $\mathcal{Q}^n$.

Next, we fix any atomic vertex $v_0 \in V(\Gamma')$. We define an auxiliary labelled rooted tree $\tau_0 = \tau_0(v_0)$ by performing a depth-first search on Γ' rooted at v_0 and then iteratively removing all leaves which are inner tree vertices. This results in a tree τ_0 rooted at an atomic vertex v_0 and all whose leaves are atomic vertices. Let $m \coloneqq |V(\tau_0)| - 1$, and let the vertices of τ_0 be labelled as $v_0, v_1, \ldots, v_m$, with the labelling given by the order in which each vertex is explored by the depth-first search performed on Γ'. For each $i \in [m]$, we define τ_i as the maximal subtree of τ_0 which contains v_i and all whose vertices have labels which are at least as large as i. Given any vertex $x \in V(I)$, we say that x is *represented* in τ_0 if $x \in V(\tau_0)$ or there exists some atomic vertex $v \in V(\tau_0)$ such that $x \in V(C(v))$. Similarly, we say that a cube $C \in \mathcal{C}$ is *represented* in τ_0 if there exists an atomic vertex $v \in V(\tau_0)$ such that $C = C(v)$. We will sometimes also say that $\mathcal{M}_x$ or $\mathcal{M}_C$ are represented in τ_0, respectively.

The tree τ_0 will be the backbone upon which we construct our long cycle. First, we need to set up some more notation. For each $i \in [m]_0$, let $p_i := d_{\tau_i}(v_i)$ and let $N_{\tau_i}(v_i) = \{u_1^i, \ldots, u_{p_i}^i\}$. It follows from (2.7.14) that

$$p_i \leq 12D - 1 \text{ if } v_i \text{ is an inner tree vertex, and } \Delta(\tau_0) \leq 12 \cdot 2^\ell D. \tag{2.7.15}$$

For each $i \in [m]_0$ and $k \in [p_i]$, let $e_k^i := \{v_i, u_k^i\}$, let $f_k^i := D(e_k^i)$, and let j_k^i be the label of u_k^i in τ_0, that is, $u_k^i = v_{j_k^i}$. For any $k \in [p_i]$, we will sometimes refer to i as the *parent index* of j_k^i. Furthermore, for each $i \in [m]_0$ such that v_i is an atomic vertex, and for each $k \in [p_i]$, consider the edge in $I(G')$ from which e_k^i originates and let ν_k^i be its endpoint in $C(v_i)$. Finally, for each $i \in [m]_0$, we define a parameter $\Delta(v_i)$ recursively by setting

$$\Delta(v_i) := \begin{cases} 0 & \text{if } v_i \text{ is an atomic vertex which is a leaf of } \tau_0, \\ \sum_{k=1}^{p_i} \Delta(u_k^i) & \text{if } v_i \text{ is an atomic vertex which is not a leaf of } \tau_0, \\ p_i + 1 + \sum_{k=1}^{p_i} \Delta(u_k^i) & \text{if } v_i \text{ is an inner tree vertex.} \end{cases} \tag{2.7.16}$$

This parameter $\Delta(v_i)$ will be used to keep track of parities throughout the following steps. Note that $\Delta(v_i)$ counts the number of times a depth first search of τ_i (starting and ending at v_i) traverses an inner tree vertex.

Consider the partition of all molecules into slices of size q introduced at the beginning of Step 3, where q is as defined in (2.7.1). Given any $v \in V(\tau_0)$, we denote the slices of its molecule by $\mathcal{M}_1(v), \ldots, \mathcal{M}_t(v)$, where t is as defined in (2.7.1). Thus, for each $i \in [t]$ we have that $\mathcal{M}_i(v) = \bigcup_{j=(i-1)q+1}^{iq} \mathcal{A}_j(v)$. For each $i \in [m]_0$, we are going to assign an *input slice* $\mathcal{M}_{b(i)}(v_i)$ to each vertex v_i. We do so by recursively assigning an *input index* $b(i) \in [t]$ to each $i \in [m]_0$. We begin by letting $b(0) := 1$. Then, for each $i \in [m]_0$ and each $k \in [p_i]$, we set

$$b(j_k^i) := \begin{cases} b(i) & \text{if } v_i \text{ is an inner tree vertex,} \\ b(i) + k - 1 \pmod t & \text{if } v_i \text{ is an atomic vertex.} \end{cases}$$

Note that the bound on $\Delta(\tau_0)$ in (2.7.15) and the definition of t in (2.7.1) imply that $b(j_k^i) \neq b(j_{k'}^i)$ whenever v_i is an atomic vertex and $k \neq k'$.

Step 11: Finding an external skeleton for T^*. Our next goal is to find an almost spanning cycle in G' by using τ_0 to explore different molecules in a given order. For this, we are going to generate a *skeleton*; this will be an ordered list of vertices which we will denote by $\mathcal{L}$. In order to construct $\mathcal{L}$, we will construct disjoint *partial skeletons* $\mathcal{L}_i$ and $\hat{\mathcal{L}}_i$ for all $i \in [m]$ in an inductive way. Each of these skeletons will start and end in the input slice for the vertex v_i which is being considered. These partial skeletons will depend on the starting and ending vertices of $\mathcal{M}_{b(i)}(v_i)$ which are provided for each of them. Therefore, given two distinct starting vertices $x, \hat{x} \in V(\mathcal{M}_{b(i)}(v_i))$ and two distinct ending vertices $y, \hat{y} \in V(\mathcal{M}_{b(i)}(v_i))$, we will denote the partial skeletons by $\mathcal{L}_i(x, y)$ and $\hat{\mathcal{L}}_i(\hat{x}, \hat{y})$, respectively.

The first step in the construction of $\mathcal{L}$ is to construct a set of vertices $L^\bullet$, to which we will refer as an *external skeleton*, and for which we will in turn construct *partial external skeletons* in an inductive way. The external skeleton will be essential in determining which vertices will not be covered by the almost spanning cycle, and hence need to be absorbed. Roughly speaking, the external skeleton will contain

(i) all vertices where the almost spanning cycle enters and leaves each cube molecule represented in τ_0, and

(ii) all vertices which are not in cube molecules and are needed to connect cube molecules to each other (that is, some clones of inner tree vertices).

On the other hand, all vertices in a vertex molecule represented in τ_0 by an inner tree vertex which do not belong to the external skeleton will have to be absorbed.

For each $i \in [m]$, given the starting and ending vertices $x, y, \hat{x}, \hat{y} \in V(\mathcal{M}_{b(i)}(v_i))$ for $\mathcal{L}_i(x, y)$ and $\hat{\mathcal{L}}_i(\hat{x}, \hat{y})$, we will denote the corresponding partial external skeleton by $L_i^\bullet(x, y, \hat{x}, \hat{y})$.

The external skeleton is constructed recursively. The partial external skeletons are the result of each recursive step, assuming that the starting and ending points have been defined. Roughly speaking, for each $i \in [m]$, we will define partial external skeletons for any possible starting and ending vertices. The starting and ending vertices which we actually use are then fixed by the partial external skeleton whose index is the parent of i. Ultimately, all of them will be fixed when defining the external skeleton $L^\bullet$.

Let $\mathcal{M}_{\mathrm{Res}} \subseteq V(\mathcal{Q}^n)$ be the union of all the clones of R'. We will construct an external skeleton $L^\bullet$ which satisfies the following properties:

(ES1) For each $i \in [m]$ such that v_i is an inner tree vertex, $L^\bullet \cap V(\mathcal{M}_{b(i)}(v_i))$ contains exactly $2p_i + 2$ vertices, half of them of each parity, and $L^\bullet \cap (V(\mathcal{M}(v_i)) \setminus V(\mathcal{M}_{b(i)}(v_i))) = \varnothing$.

(ES2) For each $i \in [m]$ such that v_i is an atomic vertex, $L^\bullet \cap V(\mathcal{M}(v_i))$ contains exactly $4p_i + 4$ vertices. If v_i is not a leaf of τ_0, eight of these vertices (four of each parity) lie in $V(\mathcal{M}_{b(i)}(v_i))$, and four (two of each parity) lie in each $V(\mathcal{M}_{b(i)+k}(v_i))$ with $k \in [p_i - 1]$. If v_i is a leaf, then all four of these vertices lie in $V(\mathcal{M}_{b(i)}(v_i))$.

(ES3) $L^\bullet \cap V(\mathcal{M}(v_0))$ contains exactly $4p_0$ vertices, four of them (two of each parity) lying in each $V(\mathcal{M}_k(v_0))$ with $k \in [p_0]$.

(ES4) The sets described in (ES1)–(ES3) partition $L^\bullet$.

(ES5) $L^\bullet \cap \mathcal{M}_{\mathrm{Res}} = \varnothing$.

We now proceed to define the partial external skeletons formally. The construction proceeds by induction on $i \in [m]$ in decreasing order, starting with $i = m$. We define a *valid connection sequence* $(x^i, y^i, \hat{x}^i, \hat{y}^i)$ *for* v_i as any set of distinct vertices $x^i, y^i, \hat{x}^i, \hat{y}^i \in V(\mathcal{M}_{b(i)}(v_i))$ which satisfy the following:

(V1) $x^i \neq_{\mathrm{p}} y^i$ if $\Delta(v_i)$ is even, and $x^i =_{\mathrm{p}} y^i$ otherwise;

(V2) $\hat{x}^i \neq_{\mathrm{p}} x^i$, and

(V3) $\hat{y}^i \neq_{\mathrm{p}} y^i$.

Given any valid connection sequence $(x^i, y^i, \hat{x}^i, \hat{y}^i)$, we will refer to x^i and $\hat{x}^i$ as *starting* vertices, and to y^i and $\hat{y}^i$ as *ending* vertices. Throughout the construction ahead, observe that, every time we use a partial external skeleton to build a larger one, its starting and ending vertices form a valid connection sequence by construction. The vertices x^i, y^i, etc. will be part of $\mathcal{L}_i(x^i, y^i)$, and the vertices $\hat{x}^i$, $\hat{y}^i$, etc. will be part of $\hat{\mathcal{L}}_i(\hat{x}^i, \hat{y}^i)$. The vertices x^i, y^i, $\hat{x}^i$, $\hat{y}^i$ will be used by the skeleton to move from the molecule represented by v_i in τ_0 to the molecule represented by its parent. Given these vertices, the following construction provides the vertices w_k^i and $\hat{w}_k^i$ (as well as z_k^i and $\hat{z}_k^i$, if applicable) which are used to move to molecules represented by the children of v_i. Given any vertices $(x, y, \hat{x}, \hat{y})$ in $\mathcal{Q}^n$ and any direction $f \in \mathcal{D}(\mathcal{Q}^n)$, we write $f + (x, y, \hat{x}, \hat{y}) = (f + x, f + y, f + \hat{x}, f + \hat{y})$.

Now suppose that $i \in [m]$ and that, for each $i' \in [m] \setminus [i]$, we have already constructed a partial external skeleton $L_{i'}^\bullet(x^{i'}, y^{i'}, \hat{x}^{i'}, \hat{y}^{i'})$ for $v_{i'}$ and every valid connection sequence $(x^{i'}, y^{i'}, \hat{x}^{i'}, \hat{y}^{i'})$ for $v_{i'}$. We will now construct a partial external skeleton for v_i and every valid connection sequence for v_i. We consider several cases.

Case 1: $v_i \in V(\tau_0)$ is a leaf of τ_0. Assume that $(x^i, y^i, \hat{x}^i, \hat{y}^i)$ is a valid connection sequence for v_i. Then, the partial external skeleton for this connection sequence is given by $L_i^\bullet(x^i, y^i, \hat{x}^i, \hat{y}^i) \coloneqq \{x^i, y^i, \hat{x}^i, \hat{y}^i\}$.

Case 2: $v_i \in V(\tau_0)$ is an inner tree vertex. We construct a set of partial external skeletons for v_i as follows.

1. Suppose $(x^i, y^i, \hat{x}^i, \hat{y}^i)$ is a valid connection sequence for v_i. Let $w_0^i \coloneqq x^i$, $w_{p_i}^i \coloneqq y^i$, $\hat{w}_0^i \coloneqq \hat{x}^i$ and $\hat{w}_{p_i}^i \coloneqq \hat{y}^i$. Let $W_0^i \coloneqq \{w_0^i, w_{p_i}^i, \hat{w}_0^i, \hat{w}_{p_i}^i\}$.

2. For each $k \in [p_i - 1]$, iteratively choose two vertices $w_k^i, \hat{w}_k^i \in V(\mathcal{M}_{b(i)}(v_i)) \setminus W_{k-1}^i$

such that $f^i_k + (w^i_{k-1}, w^i_k, \hat{w}^i_{k-1}, \hat{w}^i_k)$ is a valid connection sequence for u^i_k, and let $W^i_k \coloneqq W^i_{k-1} \cup \{w^i_k, \hat{w}^i_k\}$.

Note that the definition of q in (2.7.1) and the bound on p_i in (2.7.15) ensure that we have sufficiently many vertices to choose from (similar comments apply in the other cases). Moreover, (2.7.16) implies that $f^i_{p_i} + (w^i_{p_i-1}, w^i_{p_i}, \hat{w}^i_{p_i-1}, \hat{w}^i_{p_i})$ is a valid connection sequence for $u^i_{p_i}$. The partial external skeleton for v_i and connection sequence $(x^i, y^i, \hat{x}^i, \hat{y}^i)$ is defined as

$$L^\bullet_i(x^i, y^i, \hat{x}^i, \hat{y}^i) \coloneqq \{x^i, \hat{x}^i\} \cup \bigcup_{k=1}^{p_i} \left(\{w^i_k, \hat{w}^i_k\} \cup L^\bullet_{j^i_k}(f^i_k + (w^i_{k-1}, w^i_k, \hat{w}^i_{k-1}, \hat{w}^i_k)) \right).$$

Case 3: $v_i \in V(\tau_0)$ is an atomic vertex which is not a leaf. We construct a set of partial external skeletons for v_i as follows.

1. Assume $(x^i, y^i, \hat{x}^i, \hat{y}^i)$ is a valid connection sequence for v_i. Let $w^i_0 \coloneqq x^i$.

2. For each $k \in [p_i]$, iteratively choose distinct vertices $z^i_k, w^i_k, \hat{z}^i_k, \hat{w}^i_k \in (V(\mathcal{M}_{b(i)+k-1}(v_i)) \cap V(\mathcal{M}_{\nu^i_k})) \setminus \{x^i, y^i, \hat{x}^i, \hat{y}^i\}$ satisfying that $z^i_k \neq_{\mathrm{p}} w^i_{k-1}$ and $f^i_k + (z^i_k, w^i_k, \hat{z}^i_k, \hat{w}^i_k)$ is a valid connection sequence for u^i_k.

Then, the partial external skeleton for v_i and connection sequence $(x^i, y^i, \hat{x}^i, \hat{y}^i)$ is defined as

$$L^\bullet_i(x^i, y^i, \hat{x}^i, \hat{y}^i) \coloneqq \{x^i, y^i, \hat{x}^i, \hat{y}^i\} \cup \bigcup_{k=1}^{p_i} \left(\{z^i_k, w^i_k, \hat{z}^i_k, \hat{w}^i_k\} \cup L^\bullet_{j^i_k}(f^i_k + (z^i_k, w^i_k, \hat{z}^i_k, \hat{w}^i_k)) \right).$$

After having constructed all these partial external skeletons for all v_i with $i \in [m]$, we are now ready to construct $L^\bullet$.

1. Choose any vertex $w^0_0 \in V(\mathcal{A}_1(v_0))$.

2. For each $k \in [p_0]$, iteratively choose four distinct vertices $z^0_k, \hat{z}^0_k, w^0_k, \hat{w}^0_k \in (V(\mathcal{M}_k(v_0)) \cap V(\mathcal{M}_{\nu^0_k}))$ satisfying that $z^0_k \neq_{\mathrm{p}} w^0_{k-1}$ and $f^0_k + (z^0_k, w^0_k, \hat{z}^0_k, \hat{w}^0_k)$ is a valid connection sequence for u^0_k.

Then, we define

$$L^\bullet \coloneqq \bigcup_{k=1}^{p_0} \left(\{z^0_k, w^0_k, \hat{z}^0_k, \hat{w}^0_k\} \cup L^\bullet_{j^0_k}(f^0_k + (z^0_k, w^0_k, \hat{z}^0_k, \hat{w}^0_k)) \right).$$

Observe that (ES1)–(ES4) hold by construction. In turn, (ES5) holds because of the definition of τ_0. Indeed, observe that $V(T^*) \cap R' = \varnothing$ by (C2). Moreover, by the construction above, all vertices in $L^\bullet$ are incident to some edge in a clone of the tree T^*, and thus, they cannot lie in $\mathcal{M}_{\text{Res}}$.

Step 12: Constructing an auxiliary tree τ_0'. In order to extend the external skeleton into the skeleton and construct an almost spanning cycle, we first need to extend τ_0 to a new auxiliary tree τ_0' which encodes information about some additional molecules.

We construct τ_0' by appending some new leaves to τ_0. Note that τ_0 was built by encoding all the information about T^*, and τ_0' will encode the information about T''. In particular, by (C2), each cube $C \in \mathcal{C}'$ which intersects T'' and does not intersect T^* contains at least one vertex u which is joined to T^* by an edge $e' = \{u, v\} \in E(T'')$ such that $v \in V(C')$, where $C \neq C' \in \mathcal{C}''$. Note that the construction of τ_0 implies that C' is represented in τ_0. For each such cube C, choose one such vertex u and append a new vertex to the atomic vertex representing C' in τ_0 via an edge e which originates as $e' \in E(T'')$. We say that this newly added vertex is atomic and *represents* C. The resulting tree after all these leaves are appended is τ_0'. In particular, $\tau_0 \subseteq \tau_0'$, and it now follows that precisely the $C \in \mathcal{C}'''$ are represented in τ_0', where $\mathcal{C}'''$ is as defined in Step 9. Furthermore, it follows from (C1) that

$$\begin{aligned} &d_{\tau_0'}(v) \leq 12D \text{ for all } v \in V(\tau_0') \text{ which are inner tree vertices, and} \\ &\Delta(\tau_0') \leq 12 \cdot 2^\ell D. \end{aligned} \tag{2.7.17}$$

For all vertices of τ_0', we will use the same notation for the vertices, cubes and molecules that they represent as we did for the vertices of τ_0. Note that, by (C4.3) and (C5.3),

(CP) every cube C belonging to some absorbing ℓ-cube pair in $\mathcal{C}^{\text{sc}} \cup \mathcal{C}^{\neg\text{sc}}$ is represented in τ_0'.

It will be important for us that τ_0' represents 'most' vertices of the hypercube. In particular, for each $x \in V(I)$, let $\lambda(x)$ denote the number of vertices $y \in N_I(x)$ which are represented in τ_0' by atomic vertices. By (C3), we have that

$$\lambda(x) \geq (1 - 2/\ell^4)n. \tag{2.7.18}$$

By an averaging argument, it follows that at least $(1 - 2/\ell^4)2^{n-s}$ vertices $x \in V(I)$ are represented in τ_0' by atomic vertices. We will construct an almost spanning cycle in G' which contains all the clones of these vertices.

Let $m' := |V(\tau_0')| - 1$. Label $V(\tau_0') \setminus V(\tau_0) = \{v_{m+1}, \dots, v_{m'}\}$ arbitrarily. For each $i \in [m]$, we define τ_i' as the maximal subtree of τ_0' which contains v_i and all of whose vertices have labels at least as large as i. For each $i \in [m]_0$, let $p_i' := d_{\tau_i'}(v_i)$ and let $N_{\tau_i'}(v_i) = \{u_1^i, \dots, u_{p_i'}^i\}$ (where the labelling is consistent with that of $N_{\tau_i}(v_i)$). For each $i \in [m]_0$ and $k \in [p_i'] \setminus [p_i]$, let $e_k^i := \{v_i, u_k^i\}$, let $f_k^i := D(e_k^i)$, and let j_k^i be the label of u_k^i in τ_0'. Furthermore, for each $i \in [m]_0$ such that v_i is an atomic vertex, and for each $k \in [p_i'] \setminus [p_i]$, consider the unique edge which e_k^i originates from in $I(G')$ and let ν_k^i be its endpoint in $C(v_i)$. Finally, for each $i \in [m'] \setminus [m]$ we set $\Delta(v_i) := 0$.

As in Step 10, we consider the partition into slices for the new molecules arising from the newly added cubes represented by τ_0'. For each $i \in [m'] \setminus [m]$, we assign an input index $b(i) \in [t]$. To do so, for each $i \in [m]_0$ such that v_i is an atomic vertex and each $k \in [p_i'] \setminus [p_i]$, we set $b(j_k^i) := b(i) + k - 1 \pmod{t}$. Similarly to Step 10, (2.7.1) and (2.7.17) imply that in this case $b(j_k^i) \neq b(j_{k'}^i)$ for all $k \neq k'$. For each $i \in [m'] \setminus [m]$, let ℓ_i be the label in τ_0' of the unique vertex adjacent to v_i (i.e., the parent label of i), and let m_i be the label of v_i in $N_{\tau_{\ell_i}'}(v_{\ell_i})$. Note that $b(i) = b(\ell_i) + m_i - 1$.

Step 13: Fixing absorbing ℓ-cube pairs for vertices that need to be absorbed. At this point, we can determine every vertex in $V(\mathcal{Q}^n)$ that will have to be absorbed into the almost spanning cycle we are going to construct. For every vertex $x \in V(I)$ not represented in τ_0', we will have to absorb all vertices in $\mathcal{M}_x$. Furthermore, for each $v \in V(\tau_0)$ which is an inner tree vertex, we will also need to absorb all vertices in $\mathcal{M}_v \setminus L^\bullet$. By (ES1), this means that, in each such molecule $\mathcal{M}_v$, the same number of vertices of each parity need to be absorbed. Recall the definition of V_{abs} from Step 9. This is precisely the set of vertices which are not represented in τ_0' by an atomic vertex and, therefore, it is the set of all vertices

$x \in V(I)$ such that some clone of x needs to be absorbed. It follows from (2.7.18) that

$$|V_{\text{abs}}| \le 2^{n-s+1}/\ell^4. \tag{2.7.19}$$

Now, for each $x \in V_{\text{abs}}$, we will pair the vertices in each slice which need to be absorbed (each pair consisting of one vertex of each parity) and fix an absorbing ℓ-cube pair for each such pair of vertices. The absorbing ℓ-cube pair that we fix will be the one given by (C4) or (C5) for this pair of vertices, depending on whether $x \in V_{\text{sc}}$ or not.

For each $x \in V_{\text{abs}}$ and $\mathcal{S} \in \mathcal{S}(\mathcal{M}_x)$, let $S(x, \mathcal{S}) := V(\mathcal{S}) \cap L^\bullet$. It follows by (ES1)–(ES4) that $|S(x, \mathcal{S})| \le 24D$ and $S(x, \mathcal{S})$ contains the same number of vertices of each parity. (Here we also use that $p_i \le 12D - 1$ for every inner tree vertex v_i by (2.7.15) and (2.7.17).) Therefore, the matching $\mathfrak{M}(\mathcal{S}, S(x, \mathcal{S}))$ defined in Step 3 is well defined. Recall that each edge $e \in \mathfrak{M}(\mathcal{S}, S(x, \mathcal{S}))$ gives rise to a unique index $i \in [K]$ via the relation $N(e) = A_i(x)$. (Here we ignore all those indices $i' \in [K]$ arising by artificially increasing the size of $\mathfrak{A}(x)$, see the beginning of Step 4.) For each $x \in V_{\text{abs}}$, let $\mathfrak{I}_x \subseteq [K]$ be the set of indices $i \in [K]$ which correspond to edges in $\bigcup_{\mathcal{S} \in \mathcal{S}(\mathcal{M}_x)} \mathfrak{M}(\mathcal{S}, S(x, \mathcal{S}))$.

For each $x \in V_{\text{abs}}$ and $i \in \mathfrak{I}_x$, as stated in (C4) and (C5), we have already fixed an absorbing ℓ-cube pair for the clones of x corresponding to (x, i). Let

$$V^{\text{abs}} := \bigcup_{x \in V_{\text{abs}}} V(\mathcal{M}_x) \setminus L^\bullet.$$

As discussed above, this is the set of all vertices that need to be absorbed. Recall that G' was defined before (C1)–(C5). It follows from (C4) and (C5) that $((H \cup G) \setminus F) \cup G'$ contains a set $\mathcal{C}^{\text{abs}} = \{(C^l(u), C^r(u)) : u \in V^{\text{abs}}\}$ of absorbing ℓ-cube pairs such that

(C_1) for all distinct $u, v \in V^{\text{abs}}$, the absorbing ℓ-cube pairs $(C^l(u), C^r(u))$ and $(C^l(v), C^r(v))$ for u and v are vertex-disjoint and $(C^l(u), C^r(u)) \cup (C^l(v), C^r(v)) - \{u, v\} \subseteq G'$;

(C_2) there exists a pairing $\mathcal{U} = \{f_1, \ldots, f_{K'}\}$ of V^{abs} such that

($C_{2.1}$) for all $i \in [K']$, if $f_i = \{u_i, u_i'\}$, then $u_i \neq_{\text{p}} u_i'$;

($C_{2.2}$) if $f_i = \{u_i, u_i'\}$, then there is a vertex $v \in V_{\rm abs}$ such that u_i and u_i' are clones of v which lie in the same slice of $\mathcal{M}_v$, and $(C^l(u_i), C^r(u_i))$ and $(C^l(u_i'), C^r(u_i'))$ are clones of the same absorbing ℓ-cube pair for v in I such that $(C^l(u_i), C^r(u_i))$ lies in the same layer as u_i and $(C^l(u_i'), C^r(u_i'))$ lies in the same layer as u_i';

($C_{2.3}$) if $u, u' \in V^{\rm abs}$ do not form a pair $f \in \mathcal{U}$, then $(C^l(u), C^r(u))$ and $(C^l(u'), C^r(u'))$ are clones of vertex-disjoint absorbing ℓ-cube pairs in I (except in the case when u, u' are clones of the same vertex $v \in V_{\rm abs}$, in which case $(C^l(u), C^r(u))$ and $(C^l(u'), C^r(u'))$ are clones of absorbing ℓ-cube pairs in I which intersect only in v);

(C_3) if we let $\mathcal{C}^* := \bigcup_{(C^l(u), C^r(u)) \in \mathcal{C}^{\rm abs}} \{C^l(u), C^r(u)\}$, then $\mathcal{C}^*$ contains either two or no clones of each cube $C \in \mathcal{C}'' \cap \mathcal{C}'''$, and every cube in $\mathcal{C}^*$ is a clone of some cube $C \in \mathcal{C}'' \cap \mathcal{C}'''$.

The pairing described in (C_2) is given by the matchings $\mathfrak{M}(\mathcal{S}, S(x, \mathcal{S}))$. Furthermore, it follows from (C4.2), (C5.2) and (ES5) that

(C_4) the set of all tips of the absorbing ℓ-cube pairs in $\mathcal{C}^{\rm abs}$ is disjoint from $L^\bullet$.

We denote by $\mathfrak{L}$, $\mathfrak{R}_1$ and $\mathfrak{R}_2$ the collections of all left absorber tips, right absorber tips, and third absorber vertices, respectively, of the absorbing ℓ-cube pairs in $\mathcal{C}^{\rm abs}$. Observe that the following properties are satisfied:

(C^*1) For all $i \in [m']_0$ such that v_i is an atomic vertex and all $j \in [t]$, we have that $|\mathfrak{L} \cap V(\mathcal{M}_j(v_i))| \in \{0, 2\}$ and, if $|\mathfrak{L} \cap V(\mathcal{M}_j(v_i))| = 2$, then these two vertices u, u' lie in different atoms of the slice and satisfy that $u \neq_{\rm p} u'$.

(C^*2) For all $i \in [m']_0$ such that v_i is an atomic vertex and all $j \in [t]$, we have that $|(\mathfrak{R}_1 \cup \mathfrak{R}_2) \cap V(\mathcal{M}_j(v_i))| \in \{0, 4\}$. If $|(\mathfrak{R}_1 \cup \mathfrak{R}_2) \cap V(\mathcal{M}_j(v_i))| = 4$, then these four vertices form two pairs such that one vertex of each pair belongs to $\mathfrak{R}_1$ and the other to $\mathfrak{R}_2$. Each of these pairs lies in a different atom of the slice and satisfies that its two vertices are adjacent in G'.

(C*3) For all $i \in [m']_0$ such that v_i is an atomic vertex and all $j \in [t]$, if $\mathfrak{L} \cap V(\mathcal{M}_j(v_i)) \neq \varnothing$, then $(\mathfrak{R}_1 \cup \mathfrak{R}_2) \cap V(\mathcal{M}_j(v_i)) = \varnothing$.

(C*4) The sets described in (C*1) and (C*2) partition $\mathfrak{L}$ and $\mathfrak{R}_1 \cup \mathfrak{R}_2$, respectively.

Indeed, (C*1)–(C*3) follow from (C_2) and (C_3), and (C*4) follows by (CP).

For each $u \in V^{\text{abs}}$, we denote the edge consisting of the right absorber tip and the third absorber vertex of $(C^l(u), C^r(u))$ by $e_{\text{abs}}(u)$, and we denote by $\mathcal{P}^{\text{abs}}(u)$ the path of length three formed by the third absorber vertex, the left absorber tip, u, and the right absorber tip, visited in this order. Note that $e_{\text{abs}}(u) \in E(G')$ by (C_1). Moreover, recall that $\mathcal{C}^{\text{abs}}$ consists of absorbing ℓ-cube pairs in $((H \cup G) \setminus F) \cup G')$. Thus, $\mathcal{P}^{\text{abs}}(u) \subseteq ((H \cup G) \setminus F) \cup G'$.

Step 14: Constructing the skeleton. We can now define the skeleton for the almost spanning cycle. Intuitively, this skeleton builds on the external skeleton by adding more structure that the cycle will have to follow. In particular, the skeleton adds the edges used to traverse from each slice in a cube molecule to its neighbouring slices, and it also incorporates the cube molecules represented in τ_0' which were not represented in τ_0. (The reason why these were not incorporated earlier is the following: if we already choose the valid connection sequences for these cube molecules in Step 12, then the tips of the absorbing cubes chosen in Step 13 might have non-empty intersection with the external skeleton, which we want to avoid, see (S7) below.) Furthermore, the skeleton gives an ordering to its vertices, and the cycle will visit the vertices of the skeleton in this order.

We will build a skeleton $\mathcal{L} = (x_1, \ldots, x_r)$, for some $r \in \mathbb{N}$, and write $\mathcal{L}^\bullet \coloneqq \{x_1, \ldots, x_r\}$. We will construct $\mathcal{L}$ in such a way that the following properties hold:

(S1) For all distinct $k, k' \in [r]$, we have that $x_k \neq x_{k'}$.

(S2) $\{x_1, x_r\} \in E(G')$.

(S3) For every $k \in [r-1]$, if x_k and x_{k+1} do not both lie in the same slice of a cube molecule represented in τ_0', then $\{x_k, x_{k+1}\} \in E(G')$. Moreover, in this case, if x_{k+1} lies in a cube molecule represented in τ_0', then x_{k+2} lies in the same slice of this cube molecule as x_{k+1}.

(S4) For every $i \in [m']_0$ and every $j \in [t]$, no three consecutive vertices of $\mathcal{L}$ lie in $\mathcal{M}_j(v_i)$ (here $\mathcal{L}$ is viewed as a cyclic sequence of vertices).

(S5) For every $i \in [m']$ such that v_i is an atomic vertex and every $j \in [t]$, we have that $|V(\mathcal{M}_j(v_i)) \cap \mathcal{L}^\bullet|$ is even and $4 \leq |V(\mathcal{M}_j(v_i)) \cap \mathcal{L}^\bullet| \leq 12$. In particular, $|V(\mathcal{M}_t(v_0)) \cap \mathcal{L}^\bullet| = 4$.

(S6) For all $k \in [r]$ except two values, we have that $x_k \neq_{\mathrm{p}} x_{k+1}$. The remaining two values $k_1, k_2 \in [r]$ correspond to two pairs of vertices $x_{k_1}, x_{k_1+1}, x_{k_2}, x_{k_2+1} \in V(\mathcal{M}_t(v_0))$. For these two values, we have that $x_{k_1} \neq_{\mathrm{p}} x_{k_2}$ and either

(i) $x_{k_1} =_{\mathrm{p}} x_{k_1+1}$ and $x_{k_2} =_{\mathrm{p}} x_{k_2+1}$, or

(ii) $x_{k_1} \neq_{\mathrm{p}} x_{k_1+1}$ and $x_{k_2} \neq_{\mathrm{p}} x_{k_2+1}$,

where $x_{k_1}, x_{k_2} \in V(\mathcal{A}_{(t-1)q+1}(v_0))$ and $x_{k_1+1}, x_{k_2+1} \in V(\mathcal{A}_{tq}(v_0))$.

(S7) $\mathcal{L}^\bullet \cap (\mathfrak{L} \cup \mathfrak{R}_1 \cup V^{\mathrm{abs}}) = \varnothing$ and $L^\bullet \subseteq \mathcal{L}^\bullet$.

As happened with the external skeleton, the skeleton is built recursively from partial skeletons, which are defined first for the leaves. This recursive construction means that the overall order in which the molecules are visited will be determined by a depth first search of the tree τ_0'. Moreover, as discussed in Section 2.2.5, for parity reasons the skeleton will actually traverse τ_0' twice. These two traversals will be 'tied together' in the final step of the construction of the skeleton.

Note that, for each $i \in [m]$, the starting and ending vertices x^i, $\hat{x}^i$, y^i, $\hat{y}^i$ for the partial skeletons for v_i are determined by the external skeleton. For each $i \in [m'] \setminus [m]$, the starting and ending vertices for the partial skeletons of v_i will be determined when constructing the partial skeleton for the parent vertex v_{ℓ_i} of v_i. In particular, when constructing the partial skeleton for v_{ℓ_i}, we will define vertices $z^{\ell_i}_{m_i}, \hat{z}^{\ell_i}_{m_i}, w^{\ell_i}_{m_i}, \hat{w}^{\ell_i}_{m_i} \in \mathcal{M}_{b(i)}(v_{\ell_i})$. Then, the starting and ending vertices for the partial skeleton of v_i will be

$$(x^i, y^i, \hat{x}^i, \hat{y}^i) \coloneqq f^{\ell_i}_{m_i} + (z^{\ell_i}_{m_i}, w^{\ell_i}_{m_i}, \hat{z}^{\ell_i}_{m_i}, \hat{w}^{\ell_i}_{m_i}). \tag{2.7.20}$$

(Recall that ℓ_i, m_i, $b(i)$ and $f^{\ell_i}_{m_i}$ were defined at the end of Step 12.)

We are now in a position to define the partial skeletons formally. The construction proceeds by induction on $i \in [m']$ in decreasing order, starting with $i = m'$. Recall from the beginning of Step 11 that, for all $i \in [m]$, $x^i, y^i \in V(\mathcal{M}_{b(i)}(v_i))$ are the starting and ending vertices for the first partial skeleton $\mathcal{L}(x^i, y^i)$ for v_i, respectively, and $\hat{x}^i, \hat{y}^i \in V(\mathcal{M}_{b(i)}(v_i))$ are the starting and ending vertices for the second partial skeleton $\hat{\mathcal{L}}(\hat{x}^i, \hat{y}^i)$ for v_i, respectively. The vertices $x^i, y^i, \hat{x}^i, \hat{y}^i$ were fixed in the construction of the external skeleton, and they form a valid connection sequence. For each $i \in [m'] \setminus [m]$, the vertices $x^i, y^i, \hat{x}^i, \hat{y}^i \in V(\mathcal{M}_{b(i)}(v_i))$ defined in (2.7.20) will also form a valid connection sequence.

Let $\mathcal{F} \coloneqq \mathfrak{L} \cup \mathfrak{R}_1 \cup L^\bullet$. For each $k \in [2^s]$, let $\hat{e}_k$ be the direction of the edges in $\mathcal{Q}^n$ between L_k and L_{k+1}. Throughout the following construction, we will often choose vertices which are used to transition between neighbouring slices, all while avoiding the set $\mathcal{F}$. Similarly to the proof of Lemma 2.7.8, all of these choices can be made by (ES2), (ES3), (C*1), (C*2), and because all cube molecules considered here are bonded in G_5 and, therefore, also in G'. (The latter holds since for each atomic vertex $v \in V(\tau'_0)$ the corresponding cube $C(v)$ satisfies $C(v) \in \mathcal{C}'$.) Whenever we mention a vertex that we do not define here, we refer to the vertex with the same notation defined when constructing the external skeleton in Step 11.

Suppose that $i \in [m']$ and that for every $i' \in [m'] \setminus ([i] \cup [m])$ and every valid connection sequence $(x^{i'}, y^{i'}, \hat{x}^{i'}, \hat{y}^{i'})$ for $v_{i'}$ we have already defined two partial skeletons $\mathcal{L}(x^{i'}, y^{i'})$, $\hat{\mathcal{L}}(\hat{x}^{i'}, \hat{y}^{i'})$ for $v_{i'}$ with this connection sequence. (As discussed above, eventually we will only use the two partial skeletons for $v_{i'}$ with connection sequence as defined in (2.7.20).)

Moreover, suppose that for every $i' \in [m] \setminus [i]$ we have already defined two partial skeletons $\mathcal{L}(x^{i'}, y^{i'})$, $\hat{\mathcal{L}}(\hat{x}^{i'}, \hat{y}^{i'})$ for $v_{i'}$ with connection sequence $(x^{i'}, y^{i'}, \hat{x}^{i'}, \hat{y}^{i'})$ (fixed by the external skeleton). If $i \in [m]$, let $(x^i, y^i, \hat{x}^i, \hat{y}^i)$ be the connection sequence for v_i fixed by the external skeleton. If $i \in [m'] \setminus [m]$, let $(x^i, y^i, \hat{x}^i, \hat{y}^i)$ be any connection sequence for v_i. We will now define the two partial skeletons for v_i with connection sequence $(x^i, y^i, \hat{x}^i, \hat{y}^i)$. We consider several cases.

Case 1: v_i is a leaf of τ_0'. We construct the partial skeletons as follows. Let $x_0^i \coloneqq x^i$ and $\hat{x}_0^i \coloneqq \hat{x}^i$. For each $k \in [t-1]_0$, iteratively choose any two vertices $y_k^i, \hat{y}_k^i \in V(\mathcal{A}_{(b(i)+k)q}(v_i)) \setminus (\mathcal{F} \cup \{x^i, y^i, \hat{x}^i, \hat{y}^i\})$ satisfying that

1. $y_k^i \neq_{\mathrm{p}} x_k^i$ and $\hat{y}_k^i \neq_{\mathrm{p}} \hat{x}_k^i$;
2. $x_{k+1}^i \coloneqq y_k^i + \hat{e}_{(b(i)+k)q} \notin \mathcal{F} \cup \{x^i, y^i, \hat{x}^i, \hat{y}^i\}$ and $\hat{x}_{k+1}^i \coloneqq \hat{y}_k^i + \hat{e}_{(b(i)+k)q} \notin \mathcal{F} \cup \{x^i, y^i, \hat{x}^i, \hat{y}^i\}$, and
3. $\{y_k^i, x_{k+1}^i\}, \{\hat{y}_k^i, \hat{x}_{k+1}^i\} \in E(G')$.

Recall that we use $\bigtimes$ to denote the concatenation of sequences. The first and second partial skeletons for v_i with connection sequence $(x^i, y^i, \hat{x}^i, \hat{y}^i)$ are given by

$$\mathcal{L}_i(x^i, y^i) \coloneqq (x^i)\left(\bigtimes_{k=0}^{t-1}(y_k^i, x_{k+1}^i)\right)(y^i) \quad \text{and} \quad \hat{\mathcal{L}}_i(\hat{x}^i, \hat{y}^i) \coloneqq (\hat{x}^i)\left(\bigtimes_{k=0}^{t-1}(\hat{y}_k^i, \hat{x}_{k+1}^i)\right)(\hat{y}^i).$$

Case 2: $v_i \in V(\tau_0)$ is an inner tree vertex. Then, the first and second partial skeletons for v_i with connection sequence $(x^i, y^i, \hat{x}^i, \hat{y}^i)$ are defined as

$$\mathcal{L}_i(x^i, y^i) \coloneqq (x^i)\bigtimes_{k=1}^{p_i}(\mathcal{L}_{j_k^i}(x^{j_k^i}, y^{j_k^i}), w_k^i) \quad \text{and} \quad \hat{\mathcal{L}}_i(\hat{x}^i, \hat{y}^i) \coloneqq (\hat{x}^i)\bigtimes_{k=1}^{p_i}(\hat{\mathcal{L}}_{j_k^i}(\hat{x}^{j_k^i}, \hat{y}^{j_k^i}), \hat{w}_k^i),$$

where j_k^i was defined in Step 10.

Case 3: $v_i \in V(\tau_0)$ is an atomic vertex which is not a leaf of τ_0'. We construct the partial skeletons for v_i as follows. (Recall that, for each $k \in [p_i'] \setminus [p_i]$, the vertex ν_k^i was defined in Step 12.)

1. For each $k \in [p_i]$, iteratively choose distinct vertices $y_k^i, \hat{y}_k^i \in V(\mathcal{A}_{(b(i)+k-1)q}(v_i)) \setminus \mathcal{F}$ such that

   1.1. $y_k^i \neq_{\mathrm{p}} w_k^i$ and $\hat{y}_k^i \neq_{\mathrm{p}} \hat{w}_k^i$;

   1.2. $x_{k+1}^i \coloneqq y_k^i + \hat{e}_{(b(i)+k-1)q} \notin \mathcal{F}$ and $\hat{x}_{k+1}^i \coloneqq \hat{y}_k^i + \hat{e}_{(b(i)+k-1)q} \notin \mathcal{F}$, and

   1.3. $\{y_k^i, x_{k+1}^i\}, \{\hat{y}_k^i, \hat{x}_{k+1}^i\} \in E(G')$.

2. If $p_i = 0$, let $x_1^i \coloneqq x^i$ and $\hat{x}_1^i \coloneqq \hat{x}^i$. For each $k \in [p_i'] \setminus [p_i]$, iteratively choose distinct vertices $z_k^i, w_k^i, \hat{z}_k^i, \hat{w}_k^i \in (V(\mathcal{M}_{b(i)+k-1}(v_i)) \cap V(\mathcal{M}_{\nu_k^i})) \setminus (\mathcal{F} \cup \{x_k^i, \hat{x}_k^i\})$ and distinct vertices $y_k^i, \hat{y}_k^i \in V(\mathcal{A}_{(b(i)+k-1)q}(v_i)) \setminus (\mathcal{F} \cup \{z_k^i, w_k^i, \hat{z}_k^i, \hat{w}_k^i\})$ satisfying that

   2.1. $z_k^i, \hat{w}_k^i \neq_{\mathrm{p}} x_k^i$ and $\hat{z}_k^i, w_k^i =_{\mathrm{p}} x_k^i$;

   2.2. $x^{j_k^i}, y^{j_k^i}, \hat{x}^{j_k^i}, \hat{y}^{j_k^i} \notin \mathcal{F}$, where $x^{j_k^i}$, $y^{j_k^i}$, $\hat{x}^{j_k^i}$ and $\hat{y}^{j_k^i}$ are defined as in (2.7.20);

   2.3. $y_k^i \neq_{\mathrm{p}} w_k^i$ and $\hat{y}_k^i \neq_{\mathrm{p}} \hat{w}_k^i$;

   2.4. $x_{k+1}^i \coloneqq y_k^i + \hat{e}_{(b(i)+k-1)q} \notin \mathcal{F}$ and $\hat{x}_{k+1}^i \coloneqq \hat{y}_k^i + \hat{e}_{(b(i)+k-1)q} \notin \mathcal{F}$, and

   2.5. $\{y_k^i, x_{k+1}^i\}, \{\hat{y}_k^i, \hat{x}_{k+1}^i\} \in E(G')$.

   As discussed earlier, observe that a choice satisfying 2.2. exists by (C*1), (C*2) and (ES2).

3. For each $k \in [t] \setminus [p_i']$, iteratively choose distinct vertices $y_k^i, \hat{y}_k^i \in V(\mathcal{A}_{(b(i)+k-1)q}(v_i)) \setminus \mathcal{F}$ satisfying that

   3.1. $y_k^i \neq_{\mathrm{p}} x_k^i$ and $\hat{y}_k^i \neq_{\mathrm{p}} \hat{x}_k^i$;

   3.2. $x_{k+1}^i \coloneqq y_k^i + \hat{e}_{(b(i)+k-1)q} \notin \mathcal{F}$ and $\hat{x}_{k+1}^i \coloneqq \hat{y}_k^i + \hat{e}_{(b(i)+k-1)q} \notin \mathcal{F}$, and

   3.3. $\{y_k^i, x_{k+1}^i\}, \{\hat{y}_k^i, \hat{x}_{k+1}^i\} \in E(G')$.

Then, we may define the first and second partial skeletons for v_i with connection sequence $(x^i, y^i, \hat{x}^i, \hat{y}^i)$ as

$$\mathcal{L}_i(x^i, y^i) \coloneqq (x^i) \left(\bigtimes_{k=1}^{p_i'} (z_k^i, \mathcal{L}_{j_k^i}(x^{j_k^i}, y^{j_k^i}), w_k^i, y_k^i, x_{k+1}^i) \right) \left(\bigtimes_{k=p_i'+1}^{t} (y_k^i, x_{k+1}^i) \right) (y^i),$$

$$\hat{\mathcal{L}}_i(\hat{x}^i, \hat{y}^i) := (\hat{x}^i) \left(\bigtimes_{k=1}^{p'_i} (\hat{z}^i_k, \hat{\mathcal{L}}_{j^i_k}(\hat{x}^{j^i_k}, \hat{y}^{j^i_k}), \hat{w}^i_k, \hat{y}^i_k, \hat{x}^i_{k+1}) \right) \left(\bigtimes_{k=p'_i+1}^{t} (\hat{y}^i_k, \hat{x}^i_{k+1}) \right) (\hat{y}^i).$$

We are now ready to construct $\mathcal{L}$. The idea is similar to that of Case 3, except that we now tie together the first and second partial skeletons in Step 1.2 below.

1. Choose any two vertices $x^0_1, \hat{x}^0_1 \in V(\mathcal{A}_1(v_0)) \setminus \mathcal{F}$ such that

   1.1. $x^0_1 =_{\mathrm{p}} w^0_0$ and $\hat{x}^0_1 \neq_{\mathrm{p}} w^0_0$;

   1.2. $y^0_t := \hat{x}^0_1 + \hat{e}_{2^s} \notin \mathcal{F}$ and $\hat{y}^0_t := x^0_1 + \hat{e}_{2^s} \notin \mathcal{F}$, and

   1.3. $\{x^0_1, \hat{y}^0_t\}, \{\hat{x}^0_1, y^0_t\} \in E(G')$.

2. For each $k \in [p_0]$, iteratively choose two distinct vertices $y^0_k, \hat{y}^0_k \in V(\mathcal{A}_{qk}(v_0)) \setminus \mathcal{F}$ such that

   2.1. $y^0_k \neq_{\mathrm{p}} w^0_k$ and $\hat{y}^0_k \neq_{\mathrm{p}} \hat{w}^0_k$;

   2.2. $x^0_{k+1} := y^0_k + \hat{e}_{kq} \notin \mathcal{F}$ and $\hat{x}^0_{k+1} := \hat{y}^0_k + \hat{e}_{kq} \notin \mathcal{F}$, and

   2.3. $\{y^0_k, x^0_{k+1}\}, \{\hat{y}^0_k, \hat{x}^0_{k+1}\} \in E(G')$.

3. For each $k \in [p'_0] \setminus [p_0]$, iteratively choose distinct vertices $z^0_k, w^0_k, \hat{z}^0_k, \hat{w}^0_k \in (V(\mathcal{M}_k(v_0)) \cap V(\mathcal{M}_{\nu^0_{k+1}})) \setminus (\mathcal{F} \cup \{x^0_k, \hat{x}^0_k\})$ and distinct vertices $y^0_k, \hat{y}^0_k \in V(\mathcal{A}_{kq}(v_0)) \setminus (\mathcal{F} \cup \{z^0_k, w^0_k, \hat{z}^0_k, \hat{w}^0_k\})$ satisfying that

   3.1. $z^0_k, \hat{w}^0_k \neq_{\mathrm{p}} x^0_k$ and $\hat{z}^0_k, w^0_k =_{\mathrm{p}} x^0_k$;

   3.2. $x^{j^0_k}, y^{j^0_k}, \hat{x}^{j^0_k}, \hat{y}^{j^0_k} \notin \mathcal{F}$, where $x^{j^0_k}$, $y^{j^0_k}$, $\hat{x}^{j^0_k}$ and $\hat{y}^{j^0_k}$ are defined as in (2.7.20);

   3.3. $y^0_k \neq_{\mathrm{p}} w^0_k$ and $\hat{y}^0_k \neq_{\mathrm{p}} \hat{w}^0_k$;

   3.4. $x^0_{k+1} := y^0_k + \hat{e}_{kq} \notin \mathcal{F}$ and $\hat{x}^0_{k+1} := \hat{y}^0_k + \hat{e}_{kq} \notin \mathcal{F}$, and

   3.5. $\{y^0_k, x^0_{k+1}\}, \{\hat{y}^0_k, \hat{x}^0_{k+1}\} \in E(G')$.

4. For each $k \in [t-1] \setminus [p'_0]$, iteratively choose any two vertices $y^0_k, \hat{y}^0_k \in V(\mathcal{A}_{kq}(v_0)) \setminus \mathcal{F}$ satisfying that

4.1. $y_k^0 \neq_{\mathrm{p}} x_k^0$ and $\hat{y}_k^0 \neq_{\mathrm{p}} \hat{x}_k^0$;

4.2. $x_{k+1}^0 \coloneqq y_k^0 + \hat{e}_{kq} \notin \mathcal{F}$ and $\hat{x}_{k+1}^0 \coloneqq \hat{y}_k^0 + \hat{e}_{kq} \notin \mathcal{F}$, and

4.3. $\{y_k^0, x_{k+1}^0\}, \{\hat{y}_k^0, \hat{x}_{k+1}^0\} \in E(G')$.

The final definition of $\mathcal{L}$ is given by

$$\mathcal{L} \coloneqq (x_1^0)\left(\bigtimes_{k=1}^{p_0'}(z_k^0, \mathcal{L}_{j_k^0}(x^{j_k^0}, y^{j_k^0}), w_k^0, y_k^0, x_{k+1}^0)\right)\left(\bigtimes_{k=p_0'+1}^{t-1}(y_k^0, x_{k+1}^0)\right)(y_t^0, \hat{x}_1^0)$$
$$\left(\bigtimes_{k=1}^{p_0'}(\hat{z}_k^0, \hat{\mathcal{L}}_{j_k^0}(\hat{x}^{j_k^0}, \hat{y}^{j_k^0}), \hat{w}_k^0, \hat{y}_k^0, \hat{x}_{k+1}^0)\right)\left(\bigtimes_{k=p_0'+1}^{t-1}(\hat{y}_k^0, \hat{x}_{k+1}^0)\right)(\hat{y}_t^0).$$

Observe that (S1)–(S6) hold by construction. In particular, (2.7.16) together with (V1) ensure that in Case 3 the final two vertices of the two partial skeletons satisfy $x_{t+1}^i \neq_{\mathrm{p}} y^i$ and $\hat{x}_{t+1}^i \neq_{\mathrm{p}} \hat{y}^i$. Moreover, the pairs x_t^0, y_t^0 and $\hat{x}_t^0, \hat{y}_t^0$ will play the roles of the pairs x_{k_1}, x_{k_1+1} and x_{k_2}, x_{k_2+1} in the second part of (S6). Similarly, (S7) holds by combining the construction of $\mathcal{L}$, (C_4), (ES5) and the definition of V^{abs}.

Recall that we write $\mathcal{L} = (x_1, \ldots, x_r)$. For each $i \in [m']_0$ such that v_i is an atomic vertex and each $j \in [t]$, let $\mathfrak{J}_{i,j} \coloneqq \{k \in [r] : x_k, x_{k+1} \in V(\mathcal{M}_j(v_i))\}$ and $S_{i,j} \coloneqq \{\{x_k, x_{k+1}\} : k \in \mathfrak{J}_{i,j}\}$.

Step 15: Constructing an almost spanning cycle. We will now apply the connecting lemmas to obtain an almost spanning cycle in G' from $\mathcal{L} = (x_1, \ldots, x_r)$. For each $i \in [m']_0$ such that v_i is an atomic vertex and each $j \in [t]$, except the pair $(0, t)$, we apply Lemma 2.7.8 to the slice $\mathcal{M}_j(v_i)$ and the graph G', with $\mathfrak{L} \cap V(\mathcal{M}_j(v_i))$, $(\mathfrak{R}_1 \cup \mathfrak{R}_2) \cap V(\mathcal{M}_j(v_i))$ and $S_{i,j}$ playing the roles of L, R and the pairs of vertices described in Lemma 2.7.8(C3), respectively. Note that the conditions of Lemma 2.7.8 can be verified as follows. (C1) and (C2) hold by (C*1) and (C*2) combined with (C*3). (C3) holds by (S1) and (S3)–(S7). For $\mathcal{M}_t(v_0)$, we apply Lemma 2.7.8 or Lemma 2.7.9 depending on whether (ii) or (i) holds in (S6) (the conditions for Lemma 2.7.9 can be checked analogously). For each $i \in [m']_0$ such that

v_i is an atomic vertex and each $j \in [t]$, this yields $|\mathfrak{J}_{i,j}|$ vertex-disjoint paths $(\mathcal{P}_k^{i,j})_{k \in \mathfrak{J}_{i,j}}$ in $\mathcal{M}_j(v_i) \cup G' = G'$ such that, for each $k \in \mathfrak{J}_{i,j}$,

(i) $\mathcal{P}_k^{i,j}$ is an (x_k, x_{k+1})-path,

(ii) $\bigcup_{k \in \mathfrak{J}_{i,j}} V(\mathcal{P}_k^{i,j}) = V(\mathcal{M}_j(v_i)) \setminus \mathfrak{L}$, and

(iii) any pair of second and third absorber vertices in $\mathfrak{R}_1 \cup \mathfrak{R}_2$ contained in the same atom of $\mathcal{M}_j(v_i)$ form an edge in one of the paths.

Now consider the path obtained as follows by going through $\mathcal{L}$. Start with x_1. For each $k \in [r]$, if there exist $i \in [m']_0$ and $j \in [t]$ such that $\{x_k, x_{k+1}\} \in S_{i,j}$, add $\mathcal{P}_k^{i,j}$ to the path; otherwise, add the edge $\{x_k, x_{k+1}\}$ (this must be an edge of G' by (S3)). Finally, add the edge $\{x_r, x_1\}$ of G' (this is given by (S2)) to the path to close it into a cycle $\mathfrak{H}$ in G'. This cycle satisfies the following properties (recall that $e_{\mathrm{abs}}(u)$ was defined at the end of Step 13):

(HC1) $|V(\mathfrak{H})| \geq (1 - 4/\ell^4)2^n$.

(HC2) $V(\mathfrak{H}) \,\dot\cup\, \mathfrak{L} \,\dot\cup\, V^{\mathrm{abs}}$ partitions $V(\mathcal{Q}^n)$.

(HC3) For all $u \in V^{\mathrm{abs}}$, we have that $e_{\mathrm{abs}}(u) \in E(\mathfrak{H})$.

Indeed, note that $\mathfrak{H}$ covers all vertices in $L^\bullet$ (since $L^\bullet \subseteq \mathcal{L}^\bullet$ by (S7)) as well as all vertices lying in cube molecules represented in τ_0' except for those in $\mathfrak{L}$ (by (ii)). Together with the definition of V^{abs}, this implies (HC2). Moreover, since $|\mathfrak{L}| = |V^{\mathrm{abs}}|$, (HC1) follows from (2.7.19). Finally, (HC3) follows by (iii).

Step 16: Absorbing vertices to form a Hamilton cycle. For each $u \in V^{\mathrm{abs}}$, replace the edge $e_{\mathrm{abs}}(u)$ by the path $\mathcal{P}_{\mathrm{abs}}(u)$ (recall from the end of Step 13 that $\mathcal{P}_{\mathrm{abs}}(u)$ lies in $((H \cup G) \setminus F) \cup G')$. Clearly, this incorporates all vertices of $\mathfrak{L} \cup V^{\mathrm{abs}}$ into the cycle and, by (HC2) and (HC3), the resulting cycle is Hamiltonian. □

2.7.5 Proofs of Theorems 2.1.1, 2.1.2 and 2.1.7

First, we show that, as a byproduct of the proof of Theorem 2.7.1, we also have a proof of Theorem 2.1.2.

Proof of Theorem 2.1.2. Apply Steps 1, 4, 6, 7, 10, 11, 12, 14 and 15 in succession. In general, any reference to absorbing cubes in these steps (see e.g. the end of Step 6) should be skipped as well. □

Next, we will show how Theorem 2.7.1 can be used to prove Theorem 2.1.7.

Proof of Theorem 2.1.7. Consider a decomposition of H into k edge-disjoint subgraphs $H_1 \cup \cdots \cup H_k$ such that, for every $i \in [k]$, we have $\delta(H_i) \geq \alpha n/(2k)$. To see that this is possible, let us randomly partition the edges of H so that each $e \in E(H)$ is assigned to one of the H_i's uniformly at random and independently from all other edges. Thus, for every $i \in [k]$ we have $\mathbb{P}[e \in E(H_i)] = 1/k$. It follows by Lemma 2.4.2 that, for every vertex $x \in V(\mathcal{Q}^n)$ and every $i \in [k]$,

$$\mathbb{P}[d_{H_i}(x) \leq \alpha n/(2k)] \leq e^{-\alpha n/(8k)}.$$

For each $x \in V(\mathcal{Q}^n)$, let $\mathcal{B}(x)$ be the event that $d_{H_i}(x) \leq \alpha n/(2k)$ for some $i \in [k]$. Hence, $\mathbb{P}[\mathcal{B}(x)] \leq ke^{-\alpha n/(8k)}$ for all $x \in V(\mathcal{Q}^n)$. Observe that $\mathcal{B}(x)$ is independent of the collection of events $\{\mathcal{B}(y) : \mathrm{dist}(x, y) \geq 2\}$. A simple application of Lemma 2.4.5 shows that

$$\mathbb{P}\left[\bigwedge_{x \in V(\mathcal{Q}^n)} \overline{\mathcal{B}(x)}\right] > 0$$

and, therefore, such a decomposition of H exists.

We now consider a similar decomposition of $\mathcal{Q}^n_\varepsilon$. In particular, given $\mathcal{Q}^n_\varepsilon$, we partition its edges into k edge-disjoint subgraphs, $Q_1 \cup \cdots \cup Q_k$, in such a way that, if $e \in E(\mathcal{Q}^n_\varepsilon)$, then e is assigned to one of the Q_i chosen uniformly at random and independently of all other edges. Thus, for each $e \in E(\mathcal{Q}^n_\varepsilon)$ we have $\mathbb{P}[e \in E(Q_i)] = 1/k$ for all $i \in [k]$. It follows that, for each $i \in [k]$, we have $Q_i \sim \mathcal{Q}^n_{\varepsilon/k}$.

Let Φ be a constant such that Theorem 2.7.1 holds with ε/k, $\alpha/(2k)$ and $k+2$ playing the roles of ε, α and c, respectively. For each $i \in [k]$, apply Theorem 2.7.1 with H_i and Q_i playing the roles of H and G, respectively. We obtain that a.a.s. there exists a subgraph $G_i \subseteq Q_i$ with $\Delta(G_i) \leq \Phi$ such that, for every $F_i \subseteq \mathcal{Q}^n$ with $\Delta(F_i) \leq (k+2)\Phi$, the graph $((H_i \cup Q_i) \setminus F_i)) \cup G_i$ is Hamiltonian. Condition on the event that this holds for all $i \in [k]$ simultaneously (which holds a.a.s. by a union bound).

We are now going to find k edge-disjoint Hamilton cycles $C_1, \dots, C_k$ iteratively. For each $i \in [k]$, we proceed as follows. Let $F_i := \bigcup_{j=1}^{k} G_j \cup \bigcup_{j=1}^{i-1} C_j$. It is clear by construction that $\Delta(F_i) \leq k(\Phi + 2) \leq (k+2)\Phi$. By the conditioning above, there must be a Hamilton cycle $C_i \subseteq ((H_i \cup Q_i) \setminus F_i)) \cup G_i$. Take any such C_i and proceed.

It remains to prove that $C_1, \dots, C_k$ are pairwise edge-disjoint. In order to see this, suppose that there exist $i, j \in [k]$ with $i < j$ such that $E(C_i) \cap E(C_j) \neq \varnothing$, and let $e \in E(C_i) \cap E(C_j)$. In order to have $e \in E(C_i)$, since $G_j \subseteq F_i \setminus G_i$, we must have $e \notin E(G_j)$. However, since $e \in F_j$ by definition, we must have $e \in E(G_j)$, a contradiction. $\square$

Now, Theorem 2.1.1 follows as an immediate corollary.

Proof of Theorem 2.1.1. It is well known (see e.g. [15]) and easy to show that $\mathcal{Q}^n_{1/2-\varepsilon}$ a.a.s. contains isolated vertices. So it suffices to consider $\mathcal{Q}^n_{1/2+\varepsilon}$ for any fixed $\varepsilon > 0$ and show that a.a.s. it contains k edge-disjoint Hamilton cycles. Let $0 < \delta \ll \varepsilon \leq 1/2$. Let $H \sim \mathcal{Q}^n_{1/2+\varepsilon/2}$ and $G \sim \mathcal{Q}^n_{\varepsilon/2}$. Note that $H \cup G \sim \mathcal{Q}^n_{\eta}$, for some $\eta \leq 1/2 + \varepsilon$. Furthermore, by Lemma 2.5.5, a.a.s. $\delta(H) \geq \delta n$. Applying Theorem 2.1.7 to $H \cup G$, we obtain the desired result. $\square$

Chapter 3

A bandwidth theorem for approximate decompositions

3.1 Introduction

Starting with Dirac's theorem on Hamilton cycles, a successful research direction in extremal combinatorics has been to find appropriate minimum degree conditions on a graph G which guarantee the existence of a copy of a (possibly spanning) graph H as a subgraph. On the other hand, several important questions and results in design theory ask for the existence of a decomposition of K_n into edge-disjoint copies of a (possibly spanning) graph H, or more generally into a suitable family of graphs $H_1, \dots, H_t$.

Here, we combine the two directions: rather than finding just a single spanning graph H in a dense graph G, we seek (approximate) decompositions of a dense regular graph G into edge-disjoint copies of spanning sparse graphs H. A specific instance of this is the recent proof of the Hamilton decomposition conjecture and the 1-factorization conjecture for large n [35]: the former states that for $r \geq \lfloor n/2 \rfloor$, every r-regular n-vertex graph G has a decomposition into Hamilton cycles and at most one perfect matching, the latter provides the corresponding threshold for decompositions into perfect matchings. In this paper, we restrict ourselves to approximate decompositions, but achieve asymptotically best possible results for a much wider class of graphs than matchings and Hamilton cycles.

3.1.1 Previous results: degree conditions for spanning subgraphs

Minimum degree conditions for spanning subgraphs have been obtained mainly for (Hamilton) cycles, trees, factors and bounded degree graphs. We now briefly discuss several of these. Recall that Dirac's theorem states that any n-vertex graph G with minimum degree at least $n/2$ contains a Hamilton cycle. More generally, Abbasi's proof [1] of the El-Zahar conjecture determines the minimum degree threshold for the existence of a copy of H in G where H is a spanning union of vertex-disjoint cycles (the threshold turns out to be $\lfloor (n + odd_H)/2 \rfloor$, where odd_H denotes the number of odd cycles in H).

Komlós, Sárközy and Szemerédi [67] proved a conjecture of Bollobás by showing that a minimum degree degree of $n/2 + o(n)$ guarantees every bounded degree n-vertex tree as a subgraph (this was later strengthened in [71, 34, 57]).

An F-*factor* in a graph G is a set of vertex-disjoint copies of F covering all vertices of G. The Hajnal–Szemerédi theorem [54] implies that the minimum degree threshold for the existence of a K_k-factor is $(1 - 1/k)n$. This was generalised to kth powers of Hamilton cycles by Komlós, Sárközy and Szemerédi [70]. The threshold for arbitrary F-factors was determined by Kühn and Osthus [78], and is given by $(1-c(F))n+O(1)$, where $c(F)$ satisfies $1/\chi(F) \leq c(F) \leq 1/(\chi(F) - 1)$ and can be determined explicitly (e.g. $c(C_5) = 2/5$, in accordance with Abbasi's result).

A far-reaching generalisation of the Hajnal–Szemerédi theorem [54] would be provided by the Bollobás–Eldridge–Catlin (BEC) conjecture. This would imply that every n-vertex graph G of minimum degree at least $(1 - 1/(\Delta + 1))n$ contains every n-vertex graph H of maximum degree at most Δ as a subgraph. Partial results include the proof for $\Delta = 3$ and large n by Csaba, Shokoufandeh and Szemerédi [36] and bounds for large Δ by Kaul, Kostochka and Yu [61].

Bollobás and Komlós conjectured that one can improve on the BEC-conjecture for graphs H with a linear structure: any n-vertex graph G with minimum degree at least $(1 - 1/k + o(1))n$ contains a copy of every n-vertex k-chromatic graph H with bounded maximum degree and small bandwidth. Here an n-vertex graph H has *bandwidth* b if there exists an ordering $v_1, \ldots, v_n$ of $V(H)$ such that all edges $v_iv_j \in E(H)$ satisfy $|i - j| \leq b$. Throughout the paper, by H being k-chromatic we mean $\chi(H) \leq k$. This conjecture was resolved by the bandwidth theorem of Böttcher, Schacht and Taraz [24]. Note that while this result is essentially best possible when considering the class of k-chromatic graphs as a whole (consider e.g. K_k-factors), the results in [1, 78] mentioned above show that there are many graphs H for which the actual threshold is significantly smaller (e.g. the C_5-factors mentioned above).

The notion of bandwidth is related to the concept of separability: An n-vertex graph H is said to be *η-separable* if there exists a set S of at most ηn vertices such that every component of $H \setminus S$ has size at most ηn. We call such a set an *η-separator* of H. In general, the notion of having small bandwidth is more restrictive than that of being separable (e.g. the n-vertex star is $1/n$-separable but has bandwidth $\lfloor n/2 \rfloor$). However, for graphs with bounded maximum degree, it turns out that these notions are actually equivalent (see [23]).

3.1.2 Previous results: (approximate) decompositions into large graphs

We say that a collection $\mathcal{H} = \{H_1, \ldots, H_s\}$ of graphs *packs* into G if there exist pairwise edge-disjoint copies of $H_1, \ldots, H_s$ in G. In cases where $\mathcal{H}$ consists of copies of a single graph H we refer to this packing as an *H-packing* in G. If $\mathcal{H}$ packs into G and $e(\mathcal{H}) = e(G)$ (where $e(\mathcal{H}) = \sum_{H \in \mathcal{H}} e(H)$), then we say that G has a *decomposition* into $\mathcal{H}$. Once again, if $\mathcal{H}$ consists of copies of a single graph H, we refer to this as an H-decomposition of G. Informally, we refer to a packing which covers almost all edges of the host graph G as an approximate decomposition.

As in the previous section, most attention so far has focussed on (Hamilton) cycles, trees, factors, and graphs of bounded degree. Indeed, a classical construction of Walecki going back to the 19th century guarantees a decomposition of K_n into Hamilton cycles whenever n is odd. As mentioned earlier, this was extended to Hamilton decompositions of regular graphs G of high degree by Csaba, Kühn, Lo, Osthus and Treglown [35] (based on the existence of Hamilton decompositions in robustly expanding graphs proved in [79]). A different generalisation of Walecki's construction is given by the Alspach problem, which asks for a decomposition of K_n into cycles of given length. This was recently resolved by Bryant, Horsley and Petterson [25].

A further famous open problem in the area is the tree packing conjecture of Gyárfás and Lehel, which says that for any collection $\mathcal{T} = \{T_1, \ldots, T_n\}$ of trees with $|V(T_i)| = i$, the complete graph K_n has a decomposition into $\mathcal{T}$. This was recently proved by Joos, Kim, Kühn and Osthus [60] for the case where n is large and each T_i has bounded degree. The crucial tool for this was the blow-up lemma for approximate decompositions of ε-regular graphs G by Kim, Kühn, Osthus and Tyomkyn [63]. In particular, this lemma implies that if $\mathcal{H}$ is a family of bounded degree n-vertex graphs with $e(\mathcal{H}) \leq (1 - o(1))\binom{n}{2}$, then K_n has an approximate decomposition into $\mathcal{H}$. This generalises earlier results of Böttcher, Hladkỳ, Piguet and Taraz [21] on tree packings, as well as results of Messuti, Rödl and Schacht [84] and Ferber, Lee and Mousset [43] on packing separable graphs. Very recently, Allen, Böttcher, Hladkỳ and Piguet [3] were able to show that one can in fact find an approximate decomposition of K_n into $\mathcal{H}$ provided that the graphs in $\mathcal{H}$ have bounded degeneracy and maximum degree $o(n/\log n)$. This implies an approximate version of the tree packing conjecture when the trees have maximum degree $o(n/\log n)$. The latter improves a bound of Ferber and Samotij [44] which follows from their work on packing (spanning) trees in random graphs.

An important type of decomposition of K_n is given by resolvable designs: a resolvable

F-design consists of a decomposition into F-factors. Ray-Chaudhuri and Wilson [88] proved the existence of resolvable K_k-designs in K_n (subject to the necessary divisibility conditions being satisfied). This was generalised to arbitrary F-designs by Dukes and Ling [38].

3.1.3 Main result: packing separable graphs of bounded degree

Our main result provides a degree condition which ensures that G has an approximate decomposition into $\mathcal{H}$ for any collection $\mathcal{H}$ of k-chromatic η-separable graphs of bounded degree. As discussed below, our degree condition is best possible in general (unless one has additional information about the graphs in $\mathcal{H}$). By the remark at the end of Section 1.1 earlier, one can replace the condition of being η-separable by that of having bandwidth at most ηn in Theorem 3.1.2. Thus our result implies a version of the bandwidth theorem of [24] in the setting of approximate decompositions.

To state our result, we first introduce the approximate K_k-decomposition threshold δ_k^{reg} for regular graphs.

Definition 3.1.1 (Approximate K_k-decomposition threshold for regular graphs)**.** *For each $k \in \mathbb{N}\backslash\{1\}$, let δ_k^{reg} be the infimum over all $\delta \geq 0$ satisfying the following: for any $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ and $r \geq \delta n$ every n-vertex r-regular graph G has a K_k-packing consisting of at least $(1-\varepsilon)e(G)/e(K_k)$ copies of K_k.*

Roughly speaking, we will pack k-chromatic graphs H into regular host graphs G of degree at least $\delta_k^{\mathrm{reg}} n$. Actually it turns out that it suffices to assume that H is 'almost' k-chromatic in the sense that H has a $(k+1)$-colouring where one colour is used only rarely. More precisely, we say that H is (k,η)-*chromatic* if there exists a proper colouring of the graph H' obtained from H by deleting all its isolated vertices with $k+1$ colours such that one of the colour classes has size at most $\eta|V(H')|$. A similar feature is also present in [24].

Theorem 3.1.2. *For all $\Delta, k \in \mathbb{N}\backslash\{1\}$, $0 < \nu < 1$ and $\max\{1/2, \delta_k^{reg}\} < \delta \leq 1$, there exist $\xi, \eta > 0$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Suppose that $\mathcal{H}$ is a* *collection of n-vertex (k,η)-chromatic η-separable graphs and G is an n-vertex graph such that*

(i) $(\delta-\xi)n \le \delta(G) \le \Delta(G) \le (\delta+\xi)n$,

(ii) $\Delta(H) \le \Delta$ *for all* $H \in \mathcal{H}$,

(iii) $e(\mathcal{H}) \le (1-\nu)e(G)$.

Then $\mathcal{H}$ packs into G.

Note that our result holds for any minor-closed family $\mathcal{H}$ of k-chromatic bounded degree graphs by the separator theorem of Alon, Seymour and Thomas [4]. Moreover, note that since $\mathcal{H}$ may consist e.g. of Hamilton cycles, the condition that G is close to regular is clearly necessary. Also, the condition $\max\{1/2, \delta_k^{\rm reg}\} < \delta$ is necessary. To see this, if $\delta_k^{\rm reg} \le 1/2$ (which holds if $k=2$), then we consider $K_{n/2-1,n/2+1}$ which does not even contain a single perfect matching, let alone an approximate decomposition into perfect matchings. If $\delta_k^{\rm reg} > 1/2$ (which holds if $k \ge 3$), then for any $\delta < \delta_k^{\rm reg}$, the definition of $\delta_k^{\rm reg}$ ensures that there exist arbitrarily large regular graphs G of degree at least δn without an approximate decomposition into copies of K_k. As a disjoint union of a single copy of K_k with $n-k$ isolated vertices satisfies (ii), this shows that the condition of $\max\{1/2, \delta_k^{\rm reg}\} < \delta$ is sharp when considering the class of all k-chromatic separable graphs (though as in the case of embedding a single copy of some H into G, it may be possible to improve the degree bound for certain families $\mathcal{H}$).

To obtain explicit estimates for $\delta_k^{\rm reg}$, we also introduce the approximate K_k-decomposition threshold δ_k^{0+} for graphs of large minimum degree.

Definition 3.1.3 (Approximate K_k-decomposition threshold)**.** *For each $k \in \mathbb{N}\backslash\{1\}$, let δ_k^{0+} be the infimum over all $\delta \ge 0$ satisfying the following: for any $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that any n-vertex graph G with $n \ge n_0$ and $\delta(G) \ge \delta n$ has a K_k-packing consisting of at least $(1-\varepsilon)e(G)/e(K_k)$ copies of K_k.*

It is easy to see that $\delta_2^{\rm reg} = \delta_2^{0+} = 0$ and $\delta_k^{\rm reg} \leq \delta_k^{0+}$. The value of δ_k^{0+} has been subject to much attention recently: one reason is that by results of [8, 49], for $k \geq 3$ the approximate decomposition threshold δ_k^{0+} is equal to the analogous threshold $\delta_k^{\rm dec}$ which ensures a 'full' K_k-decomposition of any n-vertex graph G with $\delta(G) \geq (\delta_k^{\rm dec} + o(1))n$ which satisfies the necessary divisibility conditions. A beautiful conjecture (due to Nash-Williams in the triangle case and Gustavsson in the general case) would imply that $\delta_k^{\rm dec} = 1 - 1/(k+1)$ for $k \geq 3$. On the other hand for $k \geq 3$, it is easy to modify a well-known construction (see Proposition 3.3.7) to show that $\delta_k^{\rm reg} \geq 1 - 1/(k+1)$. Thus the conjecture would imply that $\delta_k^{\rm reg} = \delta_k^{0+} = \delta_k^{\rm dec} = 1 - 1/(k+1)$ for $k \geq 3$. A result of Dross [37] implies that $\delta_3^{0+} \leq 9/10$, and a very recent result of Montgomery [85] implies that $\delta_k^{0+} \leq 1 - 1/(100k)$ (see Lemma 3.3.10). With these bounds, the following corollary is immediate.

Corollary 3.1.4. *For all $\Delta, k \in \mathbb{N}\backslash\{1\}$ and $0 < \nu, \delta < 1$, there exist $\xi > 0$ and $n_0 \in \mathbb{N}$ such that for $n \geq n_0$ the following holds for every n-vertex graph G with*

$$(\delta - \xi)n \leq \delta(G) \leq \Delta(G) \leq (\delta + \xi)n.$$

(i) Let $\mathcal{T}$ be a collection of trees such that for all $T \in \mathcal{T}$ we have $|T| \leq n$ and $\Delta(T) \leq \Delta$. Further suppose $\delta > 1/2$ and $e(\mathcal{T}) \leq (1-\nu)e(G)$. Then $\mathcal{T}$ packs into G.

(ii) Let F be an n-vertex graph consisting of a union of vertex-disjoint cycles and let $\mathcal{F}$ be a collection of copies of F. Further suppose $\delta > 9/10$ and $e(\mathcal{F}) \leq (1-\nu)e(G)$. Then $\mathcal{F}$ packs into G.

(iii) Let $\mathcal{C}$ be a collection of cycles, each on at most n vertices. Further suppose $\delta > 9/10$ and $e(\mathcal{C}) \leq (1-\nu)e(G)$. Then $\mathcal{C}$ packs into G.

(iv) Let n be divisible by k and let $\mathcal{K}$ be a collection of n-vertex K_k-factors. Further suppose $\delta > 1 - 1/(100k)$ and $e(\mathcal{K}) \leq (1-\nu)e(G)$. Then $\mathcal{K}$ packs into G.

Note that (i) can be viewed as an approximate version of the tree packing conjecture in the setting of dense (almost) regular graphs. In a similar sense, (ii) relates to the Oberwolfach conjecture, (iii) relates to the Alspach problem and (iv) relates to the existence of resolvable designs in graphs.

Moreover, the feature that Theorem 3.1.2 allows us to efficiently pack (k, η)-chromatic graphs (rather than k-chromatic graphs) gives several additional consequences, for example: if the cycles of F in (ii) are all sufficiently long, then we can replace the condition '$\delta > 9/10$' by '$\delta > 1/2$'.

If we drop the assumption of being G close to regular, then one can still ask for the size of the largest packing of bounded degree separable graphs. For example, it was shown in [35] that every sufficiently large graph G with $\delta(G) \geq n/2$ contains at least $(n-2)/8$ edge-disjoint Hamilton cycles. The following result gives an approximate answer to the above question in the case when $\mathcal{H}$ consists of (almost) bipartite graphs.

Theorem 3.1.5. *For all $\Delta \in \mathbb{N}$, $1/2 < \delta \leq 1$ and $\nu > 0$, there exist $\eta > 0$ and $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ the following holds. Suppose that $\mathcal{H}$ is a collection of n-vertex $(2, \eta)$-chromatic η-separable graphs and G is an n-vertex graph such that*

(i) $\delta(G) \geq \delta n$,

(ii) $\Delta(H) \leq \Delta$ *for all* $H \in \mathcal{H}$,

(iii) $e(\mathcal{H}) \leq \frac{(\delta+\sqrt{2\delta-1}-\nu)n^2}{4}$.

Then $\mathcal{H}$ packs into G.

The result in general cannot be improved: Indeed, for $\delta > 1/2$ the number of edges of the densest regular spanning subgraph of G is close to $(\delta + \sqrt{2\delta - 1})n^2/4$ (see [30]). So the bound in (iii) is asymptotically optimal e.g. if n is even and $\mathcal{H}$ consists of Hamilton cycles. We discuss the very minor modifications to the proof of Theorem 3.1.2 which give Theorem 3.1.5 at the end of Section 3.6.

We raise the following open questions:

- We conjecture that the error term $\nu e(G)$ in condition (iii) of Theorem 3.1.2 can be improved. Note that it cannot be completely removed unless one assumes some divisibility conditions on G. However, even additional divisibility conditions will not always ensure a 'full' decomposition under the current degree conditions: indeed, for C_4, the minimum degree threshold which guarantees a C_4-decomposition of a graph G is close to $2n/3$, and the extremal example is close to regular (see [8] for details, more generally, the decomposition threshold of an arbitrary bipartite graph is determined in [49]).

- It would be interesting to know whether the condition on separability can be omitted. Note however, that if we do not assume separability, then the degree condition may need to be strengthened.

- It would be interesting to know whether one can relax the maximum degree condition in assumption (ii) of Theorem 3.1.2, e.g. for the class of trees.

- Given the recent progress on the existence of decompositions and designs in the hypergraph setting and the corresponding minimum degree thresholds [62, 51, 50], it would be interesting to generalise (some of) the above results to hypergraphs.

Our main tool in the proof of Theorem 3.1.2 will be the recent blow-up lemma for approximate decompositions by Kim, Kühn, Osthus and Tyomkyn [63]: roughly speaking, given a set $\mathcal{H}$ of n-vertex bounded degree graphs and an n-vertex graph G with $e(\mathcal{H}) \leq (1 - o(1))e(G)$ consisting of super-regular pairs, it guarantees a packing of $\mathcal{H}$ in G (such super-regular pairs arise from applications of Szemerédi's regularity lemma). Theorem 3.3.15 gives the precise statement of the special case that we shall apply (note that the original blow-up lemma of Komlós, Sárközy and Szemerédi [68] corresponds to the case where $\mathcal{H}$ consists of a single graph).

Subsequently, Theorem 3.1.2 has been used as a key tool in the resolution of the

Oberwolfach problem in [48]. This was posed by Ringel in 1967, given an n-vertex graph H consisting of vertex-disjoint cycles, it asks for a decomposition of K_n into copies of H (if n is odd). In fact, the results in [48] go considerably beyond the setting of the Oberwolfach problem, and imply e.g. a positive resolution also to the Hamilton-Waterloo problem.

3.2 Outline of the argument

Consider a given collection $\mathcal{H}$ of k-chromatic η-separable graphs with bounded degree and a given almost-regular graph G as in Theorem 3.1.2. We wish to pack $\mathcal{H}$ into G. The approach will be to decompose G into a bounded number of highly structured subgraphs G_t and partition $\mathcal{H}$ into a bounded number of collections $\mathcal{H}_t$. We then aim to pack each $\mathcal{H}_t$ into G_t. As described below, for each $H \in \mathcal{H}_t$, most of the edges will be embedded via the blow-up lemma for approximate decompositions proved in [63].

As a preliminary step, we first apply Szemerédi's regularity lemma (Lemma 3.3.5) to G to obtain a reduced multigraph R which is almost regular. Here each edge e of R corresponds to a bipartite ε-regular subgraph of G and the density of these subgraphs does not depend on e. We can then apply a result of Pippenger and Spencer on the chromatic index of regular hypergraphs and the definition of δ_k^{reg} to find an approximate decomposition of the reduced multigraph R into almost K_k-factors. More precisely, we find a set of edge-disjoint copies of almost K_k-factors covering almost all edges of R, where an almost K_k-factor is a set of vertex-disjoint copies of K_k covering almost all vertices of R. This approximate decomposition translates into the existence of an approximate decomposition of G into '(almost-)K_k-factor blow-ups'. Here a K_k-factor blow-up consists of a bounded number of clusters $V_1, \dots, V_{kr}$ where each pair (V_i, V_j) with $\lfloor (i-1)/k \rfloor = \lfloor (j-1)/k \rfloor$ is ε-regular of density d, and crucially d does not depend on i, j. We wish to use the blow-up lemma for approximate decompositions (Theorem 3.3.15) to pack graphs into each K_k-factor blow-up. Ideally, we would like to split $\mathcal{H}$ into a bounded number of subcollections $\mathcal{H}_{t,s}$ and pack each $\mathcal{H}_{t,s}$ into a separate K_k-factor blow-up $G_{t,s}$, where the $G_{t,s} \subseteq G$ are all edge-disjoint.

There are several obstacles to this approach. The first obstacle is that (i) the K_k-factor blow-ups $G_{t,s}$ are not spanning. In particular, they do not contain the vertices in the exceptional set V_0 produced by the regularity lemma. On the other hand, if we aim to embed an n-vertex graph $H \in \mathcal{H}$ into G, we must embed some vertices of H into V_0. However, Theorem 3.3.15 does not produce an embedding into vertices outside the K_k-factor blow-up. The second obstacle is that (ii) the K_k-factor blow-ups are not connected, whereas H may certainly be (highly) connected. This is one significant difference to [24], where the existence of a structure similar to a blown-up power of a Hamilton path in R could be utilised for the embedding. A third issue is that (iii) any resolution of (i) and (ii) needs to result in a 'balanced' packing of the $H \in \mathcal{H}$, i.e. the condition $e(\mathcal{H}) \leq (1-\nu)e(G)$ means that for most $x \in V(G)$ almost all their incident edges need to be covered.

To overcome the first issue, we use the fact that H is η-separable to choose a small separating set S for H and consider the small components of $H - S$. To be able to embed (most of) H into the K_k-factor blow-up, we need to add further edges to each K_k-factor blow-up so that the resulting 'augmented K_k-factor blow-ups' have strong connectivity properties. For this, we partition $V(G)\backslash V_0$ into T disjoint 'reservoirs' $Res_1, \dots, Res_T$, where $1/T \ll 1$. We will later embed some vertices of H into V_0 using the edges between Res_t and V_0 (see Lemma 3.4.1). Here we have to embed a vertex of H onto $v \in V_0$ using only edges between v and Res_t because we do not have any control on the edges between v and a regularity cluster V_i. We explain the reason for choosing a partition into many reservoir sets (rather than choosing a single small reservoir) below.

We also decompose most of G into graphs $G_{t,s}$ so that each $G_{t,s}$ has vertex set $V(G)\backslash(Res_t \cup V_0)$ and is a K_k-factor blow-up. We then find sparse bipartite graphs $F_{t,s} \subseteq G$ connecting Res_t with $G_{t,s}$, bipartite graphs $F'_t \subseteq G$ connecting Res_t with V_0 as well as sparse graphs $G^*_t \subseteq G$ which provide connectivity within Res_t as well as between Res_t and $G_{t,s}$. The fact that $G_{t,s}$ and $G_{t,s'}$ share the same reservoir for $s \neq s'$ permits us to choose the reservoir

Res_t to be significantly larger than V_0. Moreover, as $\bigcup Res_t$ covers all vertices in $V \setminus V_0$, if the graphs F'_t are appropriately chosen, then almost all edges incident to the vertices in V_0 are available to be used at some stage of the packing process. Our aim is to pack each $\mathcal{H}_{t,s}$ into the 'augmented' K_k-factor blow-up $G_{t,s} \cup F_{t,s} \cup F'_t \cup G^*_t$. To ensure that the resulting packings can be combined into a packing of all of the graphs in $\mathcal{H}$, we will use the fact that the graphs $G_t := \bigcup_s (G_{t,s} \cup F_{t,s}) \cup F'_t \cup G^*_t$ referred to in the first paragraph are edge-disjoint for different t.

We now discuss how to find this packing of $\mathcal{H}_{t,s}$. Consider some $H \in \mathcal{H}_{t,s}$. We first use the fact that H is separable to find a partition of H which reflects the structure of (the augmentation of) $G_{t,s}$ (see Section 3.4). Then we construct an appropriate embedding ϕ_* of parts of each graph $H \in \mathcal{H}_{t,s}$ into $Res_t \cup V_0$ which covers all vertices in $Res_t \cup V_0$ (this makes crucial use of the fact that Res_t is much larger than V_0). Later we aim to use the blow-up lemma for approximate decompositions (Theorem 3.3.15) to find an embedding ϕ of the remaining vertices of H into $V(G) \backslash (Res_t \cup V_0)$. When we apply Theorem 3.3.15, we use its additional features: in particular, the ability to prescribe appropriate 'target sets' for some of the vertices of H, to guarantee the consistency between the two embeddings ϕ_* and ϕ.

An important advantage of the reservoir partition which helps us to overcome obstacle (iii) is the following: the blow-up lemma for approximate decompositions can achieve a near optimal packing, i.e. it uses up almost all available edges. This is far from being the case for the part of the embeddings that use $F_{t,s}$, F'_t and G^*_t to embed vertices into $Res_t \cup V_0$, where the edge usage might be comparatively 'imbalanced' and 'inefficient'. (In fact, we will try to avoid using these edges as much as possible in order to preserve the connectivity properties of these graphs. We will use probabilistic allocations to avoid over-using any parts of $F_{t,s}$, F'_t and G^*_t.) However, since every vertex in $V(G_0) \backslash V_0$ is a reservoir vertex for only a small proportion of the embeddings, the resulting effect of these imbalances on the overall leftover degree of the vertices in $V(G_0) \backslash V_0$ is negligible. For V_0, we will be able to assign only low degree vertices of each H to ensure that there will always be edges of F'_t available to embed their incident edges (so the overall leftover degree of the vertices in V_0 may be large).

The above discussion motivates why we use many reservoir sets which cover all vertices in $V(G)\setminus V_0$, rather than using only one vertex set Res_1 for all $H \in \mathcal{H}$. Indeed, if some vertices of G only perform the role of reservoir vertices, this might result in an imbalance of the usage of edges incident to these vertices: some vertices in the reservoir might lose incident edges much faster or slower than the vertices in the regularity clusters. Apart from the fact that a fast loss of the edges incident to one vertex can prevent us from embedding any further spanning graphs into G, a large loss of the edges incident to the reservoir is also problematic in its own right. Indeed, since we are forced to use the edges incident to the reservoir in order to be able to embed some vertices onto vertices in V_0, this would prevent us from packing any further graphs.

Another issue is that the regularity lemma only gives us ε-regular K_k-factor blow-ups while we need super-regular K_k-factor blow-ups in order to use Theorem 3.3.15. To overcome this issue, we will make appropriate adjustments to each ε-regular K_k-factor blow-up. This means that the exceptional set V_0 will actually be different for each pair t, s of indices. We can however use probabilistic arguments to ensure that this does not significantly affect the overall 'balance' of the packing. In particular, for simplicity, in the above proof sketch we have ignored this issue.

The paper is organised as follows. We collect some basic tools in Section 3.3, and we prove a lemma which finds a suitable partition of each graph $H \in \mathcal{H}$ in Section 3.4 (Lemma 3.4.1). We prove our main lemma (Lemma 3.5.1) in Section 3.5. This lemma guarantees that we can find a suitable packing of an appropriate collection $\mathcal{H}_{t,s}$ of k-chromatic η-separable graphs with bounded degree into a graph consisting of a super-regular K_k-factor blow-up $G_{t,s}$ and suitable connection graphs $F_{t,s}$, F'_t and G^*_t. In Section 3.6, we will partition G and $\mathcal{H}$ as described above. Then we will repeatedly apply Lemma 3.5.1 to construct a packing of $\mathcal{H}$ into G.

3.3 Preliminaries

3.3.1 Notation

We write $[t] := \{1, \dots, t\}$. We often treat large numbers as integers whenever this does not affect the argument. The constants in the hierarchies used to state our results are chosen from right to left. That is, if we claim that a result holds for $0 < 1/n \ll a \ll b \leq 1$, we mean there exist non-decreasing functions $f : (0, 1] \to (0, 1]$ and $g : (0, 1] \to (0, 1]$ such that the result holds for all $0 \leq a, b \leq 1$ and all $n \in \mathbb{N}$ with $a \leq f(b)$ and $1/n \leq g(a)$. We will not calculate these functions explicitly.

We use the word *graphs* to refer to simple undirected finite graphs, and refer to *multi-graphs* as graphs with potentially parallel edges, but without loops. *Multi-hypergraphs* refer to (not necessarily uniform) hypergraphs with potentially parallel edges. A *k-graph* is a k-uniform hypergraph. A *multi-k-graph* is a k-uniform hypergraph with potentially parallel edges. For a multi-hypergraph H and a non-empty set $Q \subseteq V(H)$, we define $\mathrm{mult}_H(Q)$ to be the number of parallel edges of H consisting of exactly the vertices in Q. We say that a multi-hypergraph has *edge-multiplicity* at most t if $\mathrm{mult}_H(Q) \leq t$ for all non-empty $Q \subseteq V(H)$. A *matching* in a multi-hypergraph H is a collection of pairwise disjoint edges of H. The *rank* of a multi-hypergraph H is the size of a largest edge.

We write $H \simeq G$ if two graphs H and G are isomorphic. For a collection $\mathcal{H}$ of graphs, we let $v(\mathcal{H}) := \sum_{H \in \mathcal{H}} |V(H)|$. We say a partition $V_1, \dots, V_k$ of a set V is an *equipartition* if $||V_i| - |V_j|| \leq 1$ for all $i, j \in [k]$. For a multi-hypergraph H and $A, B \subseteq V(H)$, we let $E_H(A, B)$ denote the set of edges in H intersecting both A and B. We define $e_H(A, B) := |E_H(A, B)|$. For $v \in V(H)$ and $A \subseteq V(H)$, we let $d_{H,A}(v) := |\{e \in E(H) : v \in e, e \backslash \{v\} \subseteq A\}|$. Let $d_H(v) := d_{H,V(H)}(v)$. For $u, v \in V(H)$, we define $c_H(u, v) := |\{e \in E(H) : \{u, v\} \subseteq e\}|$. Let $\Delta(H) = \max\{d_H(v) : v \in V(H)\}$ and $\delta(H) := \min\{d_H(v) : v \in V(H)\}$.

For a graph G and sets $X, A \subseteq V(G)$, we define

$$N_{G,A}(X) := \{w \in A : uw \in E(G) \text{ for all } u \in X\} \text{ and } N_G(X) := N_{G,V(G)}(X).$$

Thus $N_G(X)$ is the common neighbourhood of X in G and $N_{G,A}(\emptyset) = A$. For a set $X \subseteq V(G)$, we define $N_G^d(X) \subseteq V(G)$ to be the set of all vertices of distance at most d from a vertex in X. In particular, $N_G^d(X) = \emptyset$ for $d < 0$. Note that $N_G(X)$ and $N_G^1(X)$ are different in general as e.g. vertices with a single edge to X are included in the latter. Moreover, note that $N_G(X) \subseteq N_G^1(X)$. We say a set $I \subseteq V(G)$ in a graph G is *k-independent* if for any two distinct vertices $u, v \in I$, the distance between u and v in G is at least k (thus a 2-independent set I is an independent set). If A, $B \subseteq V(G)$ are disjoint, we write $G[A, B]$ for the bipartite subgraph of G with vertex classes A, B and edge set $E_G(A, B)$.

For two functions $\phi : A \to B$ and $\phi' : A' \to B'$ with $A \cap A' = \emptyset$, we let $\phi \cup \phi'$ be the function from $A \cup A'$ to $B \cup B'$ such that for each $x \in A \cup A'$,

$$(\phi \cup \phi')(x) := \begin{cases} \phi(x) & \text{if } x \in A, \\ \phi'(x) & \text{if } x \in A'. \end{cases}$$

For graphs H and R with $V(R) \subseteq [r]$ and an ordered partition $(X_1, \dots, X_r)$ of $V(H)$, we say that H *admits the vertex partition* $(R, X_1, \dots, X_r)$, if $H[X_i]$ is empty for all $i \in [r]$, and for any $i, j \in [r]$ with $i \neq j$ we have that $e_H(X_i, X_j) > 0$ implies $ij \in E(R)$. We say that H is *internally q-regular with respect to* $(R, X_1, \dots, X_r)$ if H admits $(R, X_1, \dots, X_r)$ and $H[X_i, X_j]$ is q-regular for each $ij \in E(R)$.

We will often use the following Chernoff bound (see e.g. Theorem A.1.16 in [5]).

Lemma 3.3.1. *[5] Suppose $X_1, \dots, X_n$ are independent random variables such that $0 \leq X_i \leq b$ for all $i \in [n]$. Let $X := X_1 + \cdots + X_n$. Then for all $t > 0$, $\mathbb{P}[|X - \mathbb{E}[X]| \geq t] \leq 2e^{-t^2/(2b^2 n)}$.*

3.3.2 Tools involving ε-regularity

In this subsection, we introduce the definitions of (ε, d)-regularity and (ε, d)-super-regularity. We then state a suitable form of the regularity lemma for our purpose. We will also state an embedding lemma (Lemma 3.3.6) which we will use later to prove our main lemma (Lemma 3.5.1).

We say that a bipartite graph G with vertex partition (A,B) is *(ε,d)-regular* if for all sets $A' \subseteq A$, $B' \subseteq B$ with $|A'| \geq \varepsilon|A|$, $|B'| \geq \varepsilon|B|$, we have $|\frac{e_G(A',B')}{|A'||B'|} - d| < \varepsilon$. Moreover, we say that G is *ε-regular* if it is (ε,d)-regular for some d. If G is (ε,d)-regular and $d_G(a) = (d \pm \varepsilon)|B|$ for $a \in A$ and $d_G(b) = (d \pm \varepsilon)|A|$ for $b \in B$, then we say G is *(ε,d)-super-regular*. We say that G is *$(\varepsilon,d)^+$-(super)-regular* if it is (ε,d')-(super)-regular for some $d' \geq d$.

For a graph R on vertex set $[r]$, and disjoint vertex subsets $V_1,\dots,V_r$ of $V(G)$, we say that G is *$(\varepsilon,d)^+$-(super)-regular with respect to the vertex partition* $(R,V_1,\dots,V_r)$ if $G[V_i,V_j]$ is $(\varepsilon,d)^+$-(super)-regular for all $ij \in E(R)$. Being (ε,d)-*(super)-regular with respect to the vertex partition* $(R,V_1,\dots,V_r)$ is defined analogously. The following observations follow directly from the definitions.

Proposition 3.3.2. *Let $0 < \varepsilon \leq \delta \leq d \leq 1$. Suppose G is an (ε,d)-regular bipartite graph with vertex partition (A,B) and let $A' \subseteq A$, $B' \subseteq B$ with $|A'|/|A|$, $|B'|/|B| \geq \delta$. Then $G[A',B']$ is $(\varepsilon/\delta,d)$-regular.*

Proposition 3.3.3. *Let $0 < \varepsilon \leq \delta \leq d \leq 1$. Suppose G is an (ε,d)-regular bipartite graph with vertex partition (A,B). If G' is a subgraph of G with $V(G') = V(G)$ and $e(G') \geq (1-\delta)e(G)$, then G' is $(\varepsilon+\delta^{1/3},d)$-regular.*

Proposition 3.3.4. *Let $0 < \varepsilon \ll d \leq 1$. Suppose G is an (ε,d)-regular bipartite graph with vertex partition (A,B). Let*

$$A' := \{a \in A : d_G(a) \neq (d \pm \varepsilon)|B|\} \text{ and } B' := \{b \in B : d_G(b) \neq (d \pm \varepsilon)|B|\}.$$

Then $|A'| \leq 2\varepsilon|A|$ and $|B'| \leq 2\varepsilon|B|$.

The next lemma is a 'degree version' of Szemerédi's regularity lemma (see e.g. [77] on how to derive it from the original version).

Lemma 3.3.5 (Szemerédi's regularity lemma). *Suppose $M, M', n \in \mathbb{N}$ with $0 < 1/n \ll 1/M \ll \varepsilon, 1/M' < 1$ and $d > 0$. Then for any n-vertex graph G, there exist a partition of $V(G)$ into $V_0, V_1, \dots, V_r$ and a spanning subgraph $G' \subseteq G$ satisfying the following.*

(i) $M' \leq r \leq M$,

(ii) $|V_0| \leq \varepsilon n$,

(iii) $|V_i| = |V_j|$ *for all* $i, j \in [r]$,

(iv) $d_{G'}(v) > d_G(v) - (d + \varepsilon)n$ *for all* $v \in V(G)$,

(v) $e(G'[V_i]) = 0$ *for all* $i \in [r]$,

(vi) for all i, j with $1 \leq i < j \leq r$, the graph $G'[V_i, V_j]$ is either empty or $(\varepsilon, d_{i,j})$-regular for some $d_{i,j} \in [d, 1]$.

The next lemma allows us to embed a small graph H into a graph G which is $(\varepsilon, d)^+$-regular with respect to a suitable vertex partition $(R, V_1, \dots, V_r)$. In our proof of Lemma 3.5.1 later on, properties $(B1)_{3.3.6}$ and $(B2)_{3.3.6}$ will help us to prescribe appropriate 'target sets' for some of the vertices when we apply the blow-up lemma for approximate decompositions (Theorem 3.3.15). There, H will be part of a larger graph that is embedded in several stages. $(B1)_{3.3.6}$ ensures that the embedding of H is compatible with constraints arising from earlier stages and $(B2)_{3.3.6}$ will ensure the existence of sufficiently large target sets when embedding vertices x in later stages (each edge of $\mathcal{M}$ corresponds to the neighbourhood of such a vertex x).

Lemma 3.3.6. *Suppose $n, \Delta \in \mathbb{N}$ with $0 < 1/n \ll \varepsilon \ll \alpha, \beta, d, 1/\Delta \leq 1$. Suppose that G, H are graphs and $\mathcal{M}$ is a multi-hypergraph on $V(H)$ with edge-multiplicity at most Δ. Suppose $V_1, \dots, V_r$ are pairwise disjoint subsets of $V(G)$ with $\beta n \leq |V_i| \leq n$ for all $i \in [r]$, and $X_1, \dots, X_r$ is a partition of $V(H)$ with $|X_i| \leq \varepsilon n$ for all $i \in [r]$. Let $f : E(\mathcal{M}) \to [r]$ be* *a function, and for all $i \in [r]$ and $x \in X_i$, let $A_x \subseteq V_i$. Let R be a graph on $[r]$. Suppose that the following hold.*

(A1)$_{3.3.6}$ *G is $(\varepsilon, d)^+$-regular with respect to $(R, V_1, \dots, V_r)$,*

(A2)$_{3.3.6}$ *H admits the vertex partition $(R, X_1, \dots, X_r)$,*

(A3)$_{3.3.6}$ *$\Delta(H) \leq \Delta$, $\Delta(\mathcal{M}) \leq \Delta$ and the rank of $\mathcal{M}$ is at most Δ,*

(A4)$_{3.3.6}$ *for all $i, j \in [r]$, if $f(e) = i$ and $e \cap X_j \neq \emptyset$, then $ij \in E(R)$,*

(A5)$_{3.3.6}$ *for all $i \in [r]$ and $x \in X_i$, we have $|A_x| \geq \alpha |V_i|$.*

Then there exists an embedding ϕ of H into G such that

(B1)$_{3.3.6}$ *for each $x \in V(H)$, we have $\phi(x) \in A_x$,*

(B2)$_{3.3.6}$ *for each $e \in \mathcal{M}$, we have $|N_G(\phi(e)) \cap V_{f(e)}| \geq (d/2)^{\Delta} |V_{f(e)}|$.*

Note that (A4)$_{3.3.6}$ implies for all $e \in E(\mathcal{M})$ that $e \cap X_{f(e)} = \emptyset$.

Proof. For each $x \in V(H)$, let $e_x := N_H(x)$ and $\mathcal{M}'$ be a multi-hypergraph on vertex set $V(H)$ with $E(\mathcal{M}') = \{e_x : x \in V(H)\}$. Since a vertex $x \in V(H)$ belongs to e_y only when $y \in N_H(x)$, we have $d_{\mathcal{M}'}(x) = d_H(x)$. So $\mathcal{M}'$ is a multi-hypergraph with rank at most Δ and $\Delta(\mathcal{M}') \leq \Delta$. Let $\mathcal{M}^* := \mathcal{M} \cup \mathcal{M}'$ and for each $e \in E(\mathcal{M}^*)$, define

$$B_e := \begin{cases} V_{f(e)} & \text{if } e \in E(\mathcal{M}), \\ A_x & \text{if } e = e_x \in E(\mathcal{M}') \text{ for } x \in V(H). \end{cases}$$

Note that by (A3)$_{3.3.6}$, we have

$$\mathcal{M}^* \text{ has rank at most } \Delta, \text{ and } \Delta(\mathcal{M}^*) \leq \Delta(\mathcal{M}) + \Delta(\mathcal{M}') \leq 2\Delta. \tag{3.3.1}$$

Let $V(H) := \{x_1, \dots, x_m\}$, and for each $i \in [m]$, we let $Z_i := \{x_1, \dots, x_i\}$. We will iteratively extend partial embeddings $\phi_0, \dots, \phi_m$ of H into G in such a way that the following hold for all $i \leq m$.

$(\Phi 1)^i_{3.3.6}$ ϕ_i embeds $H[Z_i]$ into G,

$(\Phi 2)^i_{3.3.6}$ $\phi_i(x_k) \in A_{x_k}$, for all $k \in [i]$,

$(\Phi 3)^i_{3.3.6}$ for all $e \in \mathcal{M}^*$, we have $|N_G(\phi_i(e \cap Z_i)) \cap B_e| \geq (d/2)^{|e \cap Z_i|}|B_e|$.

Note that $(\Phi 1)^0_{3.3.6}$–$(\Phi 3)^0_{3.3.6}$ hold for an empty embedding $\phi_0 : \emptyset \to \emptyset$. Assume that for some $i \in [m]$, we have already defined an embedding ϕ_{i-1} satisfying $(\Phi 1)^{i-1}_{3.3.6}$–$(\Phi 3)^{i-1}_{3.3.6}$. We will construct ϕ_i by choosing an appropriate image for x_i. Let $s \in [r]$ be such that $x_i \in X_s$, and let $S := N_G(\phi_{i-1}(Z_i \cap e_{x_i})) \cap B_{e_{x_i}}$. Thus $S \subseteq V_s$. Since $Z_{i-1} \cap e_{x_i} = Z_i \cap e_{x_i}$, we have that $(\Phi 3)^{i-1}_{3.3.6}$ implies

$$|S| \geq (d/2)^{|Z_i \cap e_{x_i}|}\alpha\beta n > (d/2)^{\Delta}\alpha\beta n > \varepsilon^{1/3}n. \tag{3.3.2}$$

For each $e \in E(\mathcal{M}^*)$ containing x_i, we consider

$$S_e := N_G(\phi_{i-1}(Z_{i-1} \cap e)) \cap B_e.$$

By $(\Phi 3)^{i-1}_{3.3.6}$, we have

$$|S_e| \geq (d/2)^{\Delta}\alpha\beta n > \varepsilon^{1/3}n. \tag{3.3.3}$$

If $e = N_H(x)$ for some $x \in X_{s'}$ with $s' \in [r]$, then we have $S_e \subseteq B_e \subseteq V_{s'}$, and $(\mathrm{A2})_{3.3.6}$ implies that $ss' \in E(R)$. Moreover, note that if $e \in \mathcal{M}$ with $f(e) = s'$ for some $s' \in [r]$, then $S_e \subseteq B_e = V_{s'}$, and $(\mathrm{A4})_{3.3.6}$ implies that $ss' \in E(R)$. Thus in any case, $(\mathrm{A1})_{3.3.6}$ implies that $G[V_s, V_{s'}]$ is (ε, d')-regular for some $d' \geq d$. Hence, Proposition 3.3.2 with (3.3.2) and (3.3.3) implies that $G[S, S_e]$ is $(\varepsilon^{1/2}, d')$-regular. Let

$$S'_e := \{v \in S : d_{G,S_e}(v) < (d/2)|S_e|\}.$$

By Proposition 3.3.4, we have $|S'_e| \leq 2\varepsilon^{1/2}n$. Thus

$$|S \setminus \bigcup_{e \in E(\mathcal{M}^*): x_i \in e} S'_e| \overset{(3.3.1)}{\geq} |S| - 2\Delta \cdot 2\varepsilon^{1/2}n \overset{(3.3.2)}{\geq} 1. \tag{3.3.4}$$

We choose $v \in S \setminus \bigcup_{e\in E(\mathcal{M}^*):x_i\in e} S'_e$, and we extend ϕ_{i-1} into ϕ_i by letting $\phi_i(x_i) := v$. Since

$$\phi_i(x_i) \in S = N_G(\phi_{i-1}(Z_i \cap e_{x_i})) \cap B_{e_{x_i}} = N_G(\phi_i(Z_i \cap N_H(x_i))) \cap A_{x_i},$$

$(\Phi1)^i_{3.3.6}$ and $(\Phi2)^i_{3.3.6}$ hold. Also, for each $e \in E(\mathcal{M}^*)$, if $x_i \notin e$, then as we have $Z_i \cap e = Z_{i-1} \cap e$,

$$|N_G(\phi_i(Z_i \cap e)) \cap B_e| = |N_G(\phi_{i-1}(Z_{i-1} \cap e)) \cap B_e| \overset{(\Phi3)^{i-1}_{3.3.6}}{\geq} (d/2)^{|Z_i\cap e|}|B_e|.$$

If $x_i \in e$, then since $\phi_i(x_i) \notin S'_e$ and $|Z_i \cap e| = |Z_{i-1} \cap e| + 1$, we have

$$|N_G(\phi_i(Z_i \cap e)) \cap B_e| \geq |N_G(\phi_i(x_i)) \cap S_e| \geq (d/2)|S_e| \overset{(\Phi3)^{i-1}_{3.3.6}}{\geq} (d/2)^{|Z_i\cap e|}|B_e|. \quad (3.3.5)$$

Thus $(\Phi3)^i_{3.3.6}$ holds. By repeating this until we have embedded all vertices of H, we obtain an embedding ϕ_m satisfying $(\Phi1)^m_{3.3.6}$–$(\Phi3)^m_{3.3.6}$. Let $\phi := \phi_m$. Then $(\Phi2)^m_{3.3.6}$ implies that $(B1)_{3.3.6}$ holds, and $(\Phi3)^m_{3.3.6}$ together with $(A3)_{3.3.6}$ and the definition of B_e implies that $(B2)_{3.3.6}$ holds. □

3.3.3 Decomposition tools

In this subsection, we first give bounds on δ_k^{reg}. The following proposition provides a lower bound for δ_k^{reg}. The proof is only a slight extension of the extremal construction given by Proposition 1.5 in [8], and thus we omit it here.

Proposition 3.3.7. *For all $k \in \mathbb{N}\setminus\{1,2\}$ we have $\delta_k^{reg} \geq 1 - 1/(k+1)$.*

It will be convenient to use that for $k \geq 2$ this lower bound implies

$$\max\{1/2, \delta_k^{\mathrm{reg}}\} \geq 1 - 1/k. \quad (3.3.6)$$

Given two graphs F and G, let $\binom{G}{F}$ denote the set of all copies of F in G. A function ψ from $\binom{G}{F}$ to $[0,1]$ is a *fractional F-packing* of G if $\sum_{F'\in\binom{G}{F}:e\in F'} \psi(F') \leq 1$ for each $e \in E(G)$ (if we have equality for each $e \in E(G)$ then this is referred to as a *fractional F-decomposition*).

Let $\nu^*_F(G)$ be the maximum value of $\sum_{F'\in\binom{G}{F}}\psi(F')$ over all fractional F-packings ψ of G. Thus $\nu^*_F(G)\le e(G)/e(F)$ and $\nu^*_F(G)=e(G)/e(F)$ if and only if G has a fractional F-decomposition. The following very recent result of Montgomery gives a degree condition which ensures a fractional K_k-decomposition in a graph.

Theorem 3.3.8. *[85] Suppose $k,n\in\mathbb{N}$ and $0<1/n\ll 1/k<1$. Then any n-vertex graph G with $\delta(G)\ge(1-1/(100k))n$ satisfies $\nu^*_{K_k}(G)=e(G)/e(K_k)$.*

The next result due to Haxell and Rödl implies that a fractional K_k-decomposition gives rise to the existence of an approximate K_k-decomposition.

Theorem 3.3.9. *[55] Suppose $n\in\mathbb{N}$ with $0<1/n\ll\varepsilon<1$. Then any n-vertex graph G has an F-packing consisting of at least $\nu^*_F(G)-\varepsilon n^2$ copies of F.*

Lemma 3.3.10. *For $k\in\mathbb{N}\backslash\{1,2\}$, we have $\delta_k^{reg}\le\delta_k^{0+}\le 1-1/(100k)$. Moreover, $\delta_2^{reg}=\delta_2^{0+}=0$ and $\delta_3^{reg}\le\delta_3^{0+}\le 9/10$.*

Proof. It is easy to see that Theorem 3.3.8 and Theorem 3.3.9 together imply that $\delta_k^{0+}\le 1-1/(100k)$. Moreover, Theorem 3.3.9 together with a result of Dross [8] implies that $\delta_3^{0+}\le 9/10$. As any graph can be decomposed into copies of K_2, we have $\delta_2^{0+}=0$. □

In the remainder of this subsection, we prove Lemma 3.3.13. In the proof of Theorem 3.1.2, we will apply it to obtain an approximate decomposition of the reduced multi-graph R into almost K_k-factors (see Section 3.6). We will use the following consequence of Tutte's r-factor theorem.

Theorem 3.3.11. *[30] Suppose $n\in\mathbb{N}$ and $0<1/n\ll\gamma\ll 1$. If G is an n-vertex graph with $\delta(G)\ge(1/2+\gamma)n$ and $\Delta(G)\le\delta(G)+\gamma^2n$, then G contains a spanning r-regular subgraph for every even r with $r\le\delta(G)-\gamma n$.*

The following powerful result of Pippenger and Spencer [86] (based on the Rödl nibble) shows that every almost regular multi-k-graph with small maximum codegree has small chromatic index.

Theorem 3.3.12. *[86] Suppose $n, k \in \mathbb{N}$ and $0 < 1/n \ll \mu \ll \varepsilon, 1/k < 1$. Suppose H is an n-vertex multi-k-graph satisfying $\delta(H) \geq (1-\mu)\Delta(H)$, and $c_H(u,v) \leq \mu\Delta(H)$ for all $u \neq v \in V(H)$. Then we can partition $E(H)$ into $(1+\varepsilon)\Delta(H)$ matchings.*

We can now combine these tools to approximately decompose an almost regular multi-graph G of sufficient degree into 'almost' K_k-factors. All vertices of G will be used in almost all these factors except the vertices in a 'bad' set V' which are not used in any factor. Moreover, the factors come in T groups of equal size such that parallel edges of G belong to different groups. As explained in Section 3.2, we will apply this to the reduced multi-graph obtained from Szemerédi's regularity lemma.

Lemma 3.3.13. *Suppose $n, k, q, T \in \mathbb{N}$ with $0 < 1/n \ll \varepsilon, \sigma, 1/T, 1/k, 1/q, \nu \leq 1/2$ and $0 < 1/n \ll \xi \ll \nu < \sigma/2 < 1$ and $\delta = \max\{1/2, \delta_k^{reg}\} + \sigma$ and q divides T. Let G be an n-vertex multi-graph with edge-multiplicity at most q, such that for all $v \in V(G)$ we have*

$$d_G(v) = (\delta \pm \xi)qn.$$

Then there exists a subset $V' \subseteq V(G)$ with $|V'| \leq \varepsilon n$ and k dividing $|V(G)\setminus V'|$, and there exist pairwise edge-disjoint subgraphs $F_{1,1}, \ldots, F_{1,\kappa}, F_{2,1}, \ldots, F_{T,\kappa}$ with $\kappa = (\delta - \nu \pm \varepsilon)\frac{qn}{T(k-1)}$ satisfying the following.

(B1)$_{3.3.13}$ *For each $(t', i) \in [T] \times [\kappa]$, we have that $V(F_{t',i}) \subseteq V(G)\setminus V'$ and $F_{t',i}$ is a vertex-disjoint union of at least $(1-\varepsilon)n/k$ copies of K_k,*

(B2)$_{3.3.13}$ *for each $v \in V(G) \setminus V'$, we have $|\{(t', i) \in [T] \times [\kappa] : v \in V(F_{t',i})\}| \geq T\kappa - \varepsilon n$,*

(B3)$_{3.3.13}$ *for all $t' \in [T]$ and $u, v \in V(G)$, we have $|\{i \in [\kappa] : u \in N_{F_{t',i}}(v)\}| \leq 1$.*

Proof. It suffices to prove the lemma for the case when $T = q$. The general case then follows by relabelling. (We can split each group obtained from the $T = q$ case into T/q equal groups arbitrarily.) We choose a new constant μ such that

$$1/n \ll \mu \ll \varepsilon, \xi, \sigma, 1/k, 1/q.$$

For an edge colouring $\phi : E(G) \to [q]$ and $c \in [q]$, we let $G^c \subseteq G$ be the subgraph with edge set $\{e \in E(G) : \phi(e) = c\}$. We wish to show that there exists an edge-colouring $\phi : E(G) \to [q]$ satisfying the following for all $v \in V(G)$ and $c \in [q]$:

$(\Phi 1)_{3.3.13}$ $d_{G^c}(v) = (\delta \pm 2\xi)n$,

$(\Phi 2)_{3.3.13}$ G^c is a simple graph.

Recall that $e_G(u,v)$ denotes the number of edges of G between u and v. For each $\{u,v\} \in \binom{V(G)}{2}$, we choose a set $A_{\{u,v\}}$ uniformly at random from $\binom{[q]}{e_G(u,v)}$. For each $e \in E(G)$, we let $\phi(e) \in [q]$ be such that ϕ is bijective between $E_G(u,v)$ and $A_{\{u,v\}}$. This ensures that $(\Phi 2)_{3.3.13}$ holds. It is easy to see that $(\Phi 1)_{3.3.13}$ also holds with high probability by using Lemma 3.3.1.

Since $\delta \geq 1/2 + \sigma$ and $\xi \ll \nu, \sigma$, Theorem 3.3.11 implies that, for each $c \in [q]$, there exists a $(\delta - \nu)n$-regular spanning subgraph G^c_* of G^c. (By adjusting ν slightly we may assume that $(\delta - \nu)n$ is an even integer.) Since $\delta - \nu > \delta_k^{\rm reg} + \sigma/2$ and $1/n \ll \mu$, the graph G^c_* has a K_k-packing $\mathcal{Q}^c := \{Q^c_1, \dots, Q^c_t\}$ of size

$$t := \frac{(\delta - \nu - \mu)n^2}{k(k-1)}. \tag{3.3.7}$$

For each $c \in [q]$, let $\mathcal{H}^c$ be the k-graph with $V(\mathcal{H}^c) = V(G^c_*)$ and $E(\mathcal{H}^c) := \{V(Q^c_i) : i \in [t]\}$. By construction of $\mathcal{H}^c$, we have

$$\Delta(\mathcal{H}^c) \leq \frac{\Delta(G^c_*)}{k-1} \leq \frac{(\delta - \nu)n}{k-1}. \tag{3.3.8}$$

As $\mathcal{Q}^c$ is a K_k-packing in G^c_*, any pair $\{u,v\} \in \binom{V(G)}{2}$ belongs to at most one edge in $\mathcal{H}^c$. Thus for $\{u,v\} \in \binom{V(G)}{2}$,

$$c_{\mathcal{H}^c}(u,v) \leq 1. \tag{3.3.9}$$

Let

$$V'' := \bigcup_{c \in [q]} \Big\{ v \in V(G) : |\{i \in [t] : v \in V(Q^c_i)\}| < \frac{1}{k-1}(\delta - \nu - \mu^{1/3})n \Big\},$$

and let V' be a set consisting of the union of V'' as well as at most $k-1$ vertices arbitrarily chosen from $V(G)\backslash V''$ such that k divides $|V(G)\backslash V'|$. Note that for each $c \in [q]$, we have

$$e(G_*^c) - e(\mathcal{Q}^c) \leq \frac{1}{2}(\delta - \nu)n^2 - \binom{k}{2} t \overset{(3.3.7)}{\leq} \mu n^2.$$

On the other hand, since G_*^c is a $(\delta-\nu)n$-regular graph, we have

$$\begin{aligned} |V'| &\leq k + 1 + \sum_{c\in[q]} \frac{1}{\mu^{1/3} n} \sum_{v \in V(G)} \left(d_{G_*^c}(v) - (k-1) d_{\mathcal{H}^c}(v)\right) \\ &= k + 1 + \sum_{c\in[q]} \frac{2(e(G_*^c) - e(\mathcal{Q}^c))}{\mu^{1/3} n} \leq \frac{3q\mu n^2}{\mu^{1/3} n} \leq \mu^{1/2} n. \end{aligned} \tag{3.3.10}$$

Let $\tilde{\mathcal{H}}^c$ be the k-graph with $V(\tilde{\mathcal{H}}^c) := V(G_*^c) \setminus V'$ and $E(\tilde{\mathcal{H}}^c) := \{e \in E(\mathcal{H}^c) : e \cap V' = \emptyset\}$. Note that for any $v \in V(\tilde{\mathcal{H}}^c) = V(\mathcal{H}^c) \setminus V'$,

$$d_{\tilde{\mathcal{H}}^c}(v) = d_{\mathcal{H}^c}(v) \pm \sum_{u \in V'} c_{\mathcal{H}^c}(u,v) \overset{(3.3.9)}{=} d_{\mathcal{H}^c}(v) \pm |V'| \overset{(3.3.10),(3.3.8)}{=} \frac{(\delta - \nu \pm 2\mu^{1/3})n}{k-1}. \tag{3.3.11}$$

Note that we obtain the final equality from the definition of V' and the assumption that $v \notin V'$. Thus for each $c \in [q]$, we have $\delta(\tilde{\mathcal{H}}^c) \geq (1-\mu^{1/4})\Delta(\tilde{\mathcal{H}}^c)$. Together with (3.3.9) and the fact that $1/n \ll \mu \ll \varepsilon, 1/k, 1/q$, this ensures that we can apply Theorem 3.3.12 to see that for each $c \in [q]$, $E(\tilde{\mathcal{H}}^c)$ can be partitioned into $\kappa' := \frac{(\delta-\nu+\varepsilon^3/q)n}{k-1}$ matchings $M_1^c, \dots, M_{\kappa'}^c$. Let

$$\mathcal{M}^c := \{M_i^c : i \in [\kappa']\} \text{ and } \mathcal{M}_*^c := \{M_i^c : i \in [\kappa'], |M_i^c| < (1-\varepsilon)n/k\}.$$

As $|M_i^c| \leq n/k$ for any $i \in [\kappa']$ and $c \in [q]$, we have

$$\frac{(\delta - \nu - 3\mu^{1/3})n^2}{k(k-1)} \overset{(3.3.10),(3.3.11)}{\leq} |E(\tilde{\mathcal{H}}^c)| = \sum_{i\in[\kappa']} |M_i^c| < \frac{|\mathcal{M}_*^c|(1-\varepsilon)n}{k} + \frac{(\kappa' - |\mathcal{M}_*^c|)n}{k}.$$

This gives

$$|\mathcal{M}_*^c| \leq \frac{(\varepsilon^3/q + 3\mu^{1/3})kn^2}{\varepsilon n k(k-1)} \leq \frac{2\varepsilon^2 n}{q(k-1)}. \tag{3.3.12}$$

We let

$$\kappa := \min_{c\in[q]}\{|\mathcal{M}^c \setminus \mathcal{M}_*^c|\} = \kappa' - \max_{c\in[q]}\{|\mathcal{M}_*^c|\} = \frac{(\delta-\nu)n \pm 2\varepsilon^2 n/q}{k-1}. \tag{3.3.13}$$

Thus, by permuting indices, we can assume that for each $c \in [q]$, we have $M_1^c, \dots, M_\kappa^c \subseteq \mathcal{M}^c \setminus \mathcal{M}_*^c$. For each $(c,i) \in [q] \times [\kappa]$, let

$$F_{c,i} := \bigcup_{j: V(Q_j^c) \in M_i^c} Q_j^c.$$

The fact that $\mathcal{M}^c \setminus \mathcal{M}_*^c$ is a collection of pairwise edge-disjoint matchings of $\tilde{\mathcal{H}}^c \subseteq \mathcal{H}^c$ together with (3.3.9) implies that, for each $c \in [q]$, the collection $\{F_{c,i} : i \in [\kappa]\}$ consists of pairwise edge-disjoint subgraphs of $G_*^c \subseteq G$, each of which is a union of at least $(1-\varepsilon)n/k$ vertex-disjoint copies of K_k. This with $(\Phi 2)_{3.3.13}$ shows that $(\mathrm{B3})_{3.3.13}$ holds. As $G_*^1, \dots, G_*^q$ are pairwise edge-disjoint subgraphs, $\{F_{c,i} : (c,i) \in [q] \times [\kappa]\}$ forms a collection of pairwise edge-disjoint subgraphs of G. Thus $(\mathrm{B1})_{3.3.13}$ holds.

Moreover, for each $c \in [q]$ and each vertex $v \in V(G) \setminus V'$, we have

$$\begin{aligned} |\{i \in [\kappa] : v \in V(F_{c,i})\}| &\geq |\{M \in \{M_1^c, \dots, M_\kappa^c\} : v \in V(M)\}| \\ &\geq |\{M \in \mathcal{M}^c : v \in V(M)\}| - (\kappa' - \kappa) \\ &\geq d_{\tilde{H}^c}(v) - \kappa' + \kappa \overset{(3.3.11)}{\geq} \kappa - \varepsilon n/q. \end{aligned}$$

Thus $(\mathrm{B2})_{3.3.13}$ holds. □

3.3.4 Graph packing tools

The following two results from [63] will allow us to pack many bounded degree graphs into appropriate super-regular blow-ups. Lemma 3.3.14 first allows us to pack graphs into internally regular graphs which still have bounded degree, and Theorem 3.3.15 allows us to pack the internally regular graphs into an appropriate dense ε-regular graph. The results in [63] are actually significantly more general, mainly because they allow for more general reduced graphs R.

Lemma 3.3.14 ([63, Lemma 7.1]). *Suppose $n, \Delta, q, s, k, r \in \mathbb{N}$ with $0 < 1/n \ll \varepsilon \ll 1/s \ll 1/\Delta, 1/k$ and $\varepsilon \ll 1/q \ll 1$ and k divides r. Suppose that $0 < \xi < 1$ is such that $s^{2/3} \leq \xi q$.*

Let R be a graph on $[r]$ consisting of r/k vertex-disjoint copies of K_k. Let $V_1, \dots, V_r$ be a partition of some vertex set V such that $|V_i| = n$ for all $i \in [r]$. Suppose for each $j \in [s]$, L_j is a graph admitting the vertex partition $(R, X_1^j, \dots, X_r^j)$ such that $\Delta(L_j) \le \Delta$ and for each $ii' \in E(R)$, we have

$$\sum_{j=1}^{s} e(L_j[X_i^j, X_{i'}^j]) = (1 - 3\xi \pm \xi)qn,$$

and $|X_i^j| \le n$. Also suppose that for all $j \in [s]$ and $i \in [r]$, we have sets $W_i^j \subseteq X_i^j$ such that $|W_i^j| \le \varepsilon n$. Then there exists a graph H on V which is internally q-regular with respect to $(R, V_1, \dots, V_r)$ and a function ϕ which packs $\{L_1, \dots, L_s\}$ into H such that $\phi(X_i^j) \subseteq V_i$, and such that for all distinct $j, j' \in [s]$ and $i \in [r]$, we have $\phi(W_i^j) \cap \phi(W_i^{j'}) = \emptyset$.

Theorem 3.3.15 (Blow-up lemma for approximate decompositions [63, Theorem 6.1])**.** *Suppose $n, q, s, k, r \in \mathbb{N}$ with $0 < 1/n \ll \varepsilon \ll \alpha, d, d_0, 1/q, 1/k \le 1$ and $1/n \ll 1/r$ and k divides r. Suppose that R is a graph on $[r]$ consisting of r/k vertex-disjoint copies of K_k. Suppose $s \le \frac{d}{q}(1 - \alpha/2)n$ and the following hold.*

(A1)$_{3.3.15}$ *G is (ε, d)-super-regular with respect to the vertex partition $(R, V_1, \dots, V_r)$.*

(A2)$_{3.3.15}$ *$\mathcal{H} = \{H_1, \dots, H_s\}$ is a collection of graphs, where each H_j is internally q-regular with respect to the vertex partition $(R, X_1, \dots, X_k)$, and $|X_i| = |V_i| = n$ for all $i \in [r]$.*

(A3)$_{3.3.15}$ *For all $j \in [s]$ and $i \in [r]$, there is a set $W_i^j \subseteq X_i$ with $|W_i^j| \le \varepsilon n$ and for each $w \in W_i^j$, there is a set $A_w^j \subseteq V_i$ with $|A_w^j| \ge d_0 n$.*

(A4)$_{3.3.15}$ *Λ is a graph with $V(\Lambda) \subseteq [s] \times \bigcup_{i=1}^{r} X_i$ and $\Delta(\Lambda) \le (1 - \alpha)d_0 n$ such that for all $(j, x) \in V(\Lambda)$ and $j' \in [s]$, we have $|\{x' : (j', x') \in N_\Lambda((j, x))\}| \le q^2$. Moreover, for all $j \in [s]$ and $i \in [r]$, we have $|\{(j, x) \in V(\Lambda) : x \in X_i\}| \le \varepsilon |X_i|$.*

Then there is a function ϕ packing $\mathcal{H}$ into G such that, writing ϕ_j for the restriction of ϕ to H_j, the following hold for all $j \in [s]$ and $i \in [r]$.

(B1)$_{3.3.15}$ $\phi_j(X_i) = V_i$,

(B2)$_{3.3.15}$ $\phi_j(w) \in A^j_w$ *for all* $w \in W^j_i$,

(B3)$_{3.3.15}$ *for all* $(j,x)(j',y) \in E(\Lambda)$, *we have that* $\phi_j(x) \neq \phi_{j'}(y)$.

3.3.5 Miscellaneous

In the proof of Theorem 3.1.2, we often partition various graphs into parts with certain properties. The next two lemmas will allow us to obtain such partitions. Lemma 3.3.16 follows by considering a random equipartition and applying concentration of the hypergeometric distribution. Lemma 3.3.17 can be proved by assigning each edge of G to $G_1, \dots, G_s$ independently at random according to $(p_1, \dots, p_s)$, and applying Lemma 3.3.1. We omit the details.

Lemma 3.3.16. *Suppose* $n, T, r \in \mathbb{N}$ *with* $0 < 1/n \ll 1/T, 1/r \leq 1$. *Let* G *be an* n*-vertex graph. Let* $V \subseteq V(G)$ *and let* $V_1 \dots, V_r$ *be a partition of* V. *Then there exists an equipartition* $Res_1, \dots, Res_T$ *of* V *such that the following hold.*

(i) For all $t \in [T]$, $i \in [r]$ *and* $v \in V(G)$, *we have* $d_{G, Res_t \cap V_i}(v) = \frac{1}{T} d_{G,V_i}(v) \pm n^{2/3}$,

(ii) for all $t \in [T]$, $i \in [r]$, *we have* $|Res_t \cap V_i| = \frac{1}{T}|V_i| \pm n^{2/3}$.

Lemma 3.3.17. *Suppose* $n, s \in \mathbb{N}$ *with* $0 < 1/n \ll \varepsilon \ll 1/s \leq 1$ *and* $m_i \in [n]$ *for each* $i \in [2]$. *Let* G *be an* n*-vertex graph. Suppose that* $\mathcal{U}$ *is a collection of* m_1 *subsets of* $V(G)$ *and* $\mathcal{U}'$ *is a collection of* m_2 *pairs of disjoint subsets of* $V(G)$ *such that each* $(U_1, U_2) \in \mathcal{U}'$ *satisfies* $|U_1|, |U_2| > n^{3/4}$. *Let* $0 \leq p_1, \dots, p_s \leq 1$ *with* $\sum_{i=1}^s p_i = 1$. *Then there exists a decomposition* $G_1, \dots, G_s$ *of* G *satisfying the following.*

(i) For all $i \in [s]$, $U \in \mathcal{U}$ *and* $v \in V(G)$, *we have* $d_{G_i,U}(v) = p_i d_{G,U}(v) \pm n^{2/3}$,

(ii) for all $i \in [s]$ *and* $(U_1, U_2) \in \mathcal{U}'$ *such that* $G[U_1, U_2]$ *is* $(\varepsilon, d_{(U_1,U_2)})$*-regular for some* $d_{(U_1,U_2)}$, *we have that* $G_i[U_1, U_2]$ *is* $(2\varepsilon, p_i d_{(U_1,U_2)})$*-regular.*

The following lemma allows us to find well-distributed subsets of a collection of large sets. The required sets can be found via a straightforward greedy approach (while avoiding the vertices which would violate (B3)$_{3.3.18}$ in each step). So we omit the details.

Lemma 3.3.18. *Suppose $n, s, r \in \mathbb{N}$ and $0 < 1/n, 1/s \ll \varepsilon \ll d < 1$. Let A be a set of size n, and for each $(i,j) \in [s] \times [r]$ let $A_{i,j} \subseteq A$ be of size at least dn, and let $m_{i,j} \in \mathbb{N} \cup \{0\}$ be such that for all $i \in [s]$ we have $\sum_{j=1}^{r} m_{i,j} \leq \varepsilon n$. Then there exist sets $B_{1,1}, \ldots, B_{s,r}$ satisfying the following.*

(B1)$_{3.3.18}$ *For all $i \in [s]$ and $j \in [r]$, we have $B_{i,j} \subseteq A_{i,j}$ with $|B_{i,j}| = m_{i,j}$,*

(B2)$_{3.3.18}$ *for all $i \in [s]$ and $j' \neq j'' \in [r]$, we have $B_{i,j'} \cap B_{i,j''} = \emptyset$,*

(B3)$_{3.3.18}$ *for all $v \in A$, we have $|\{(i,j) \in [s] \times [r] : v \in B_{i,j}\}| \leq \varepsilon^{1/2} s$.*

The following lemma guarantees a set of k-cliques in a graph G which cover every vertex a prescribed number of times.

Lemma 3.3.19. *Let $n, m, k, t \in \mathbb{N}$ and $0 < 1/n \ll 1/t \ll \sigma, 1/k < 1$ with $k \mid n$. Let G be an n-vertex graph with $\delta(G) \geq (1 - \frac{1}{k} + \sigma)n$. Suppose that for each $v \in V(G)$, we have $d_v \in [m] \cup \{0\}$. Then there exists a multi-k-graph H on vertex set $V(G)$ satisfying the following.*

(B1)$_{3.3.19}$ *For each $e \in E(H)$, we have $G[e] \simeq K_k$,*

(B2)$_{3.3.19}$ *for each $v \in V(G)$, we have $d_H(v) - d_v = (t+1)m \pm 1$.*

Proof. Let
$$m' := \max_{u,v \in V(G)} \{d_u - d_v\}.$$
Then $m' \in [m]$. For a multi-hypergraph H on vertex set $V(G)$ and $v \in V(G)$, let $p_H(v) := d_H(v) - d_v$. We will prove that for each $\ell \in [m'-1] \cup \{0\}$, there exists a hypergraph H_ℓ satisfying the following.

(H1)$^{\ell}_{3.3.19}$ For each $e \in E(H)$, we have $G[e] \simeq K_k$,

(H2)$^{\ell}_{3.3.19}$ $\Delta(H_\ell) \leq \ell(t+1)$,

(H3)$^{\ell}_{3.3.19}$ $\max_{u,v\in V(G)}\{p_{H_\ell}(v) - p_{H_\ell}(u)\} \leq m' - \ell$.

Note that $H_0 = \emptyset$ satisfies (H1)$^{0}_{3.3.19}$–(H3)$^{0}_{3.3.19}$. Assume that for some $\ell \in [m'-2] \cup \{0\}$, we have already constructed H_ℓ satisfying (H1)$^{\ell}_{3.3.19}$–(H3)$^{\ell}_{3.3.19}$. We will now construct $H_{\ell+1}$.

If $\max_{u\in V(G)}\{p_{H_\ell}(u)\} - \min_{u\in V(G)}\{p_{H_\ell}(u)\} \leq 1$, then as $\ell \leq m' - 2$, we can let $H_{\ell+1} := H_\ell$, then (H1)$^{\ell+1}_{3.3.19}$–(H3)$^{\ell+1}_{3.3.19}$ hold. Thus assume that

$$\max_{u\in V(G)}\{p_{H_\ell}(u)\} - \min_{u\in V(G)}\{p_{H_\ell}(u)\} \geq 2. \tag{3.3.14}$$

Let

$$A := \{v \in V(G) : p_{H_\ell}(v) > \min_{u\in V(G)}\{p_{H_\ell}(u)\}\} \text{ and } A_{\max} := \{v \in V(G) : p_{H_\ell}(v) = \max_{u\in V(G)}\{p_{H_\ell}(u)\}\}.$$

First assume that $|A| \geq k$. Let $A' \subseteq A$ be a set of at most $k-1$ vertices such that k divides $|A| + |A'|$ and $p_{H_\ell}(v) \geq \max_{u\in A\setminus A'} p_{H_\ell}(u)$ for all $v \in A'$. Note that we have either $A' \subseteq A_{\max}$ or $A_{\max} \subseteq A'$. Then we can take a collection $\mathcal{A} := \{A_1, \dots, A_{t+1}\}$ of (possibly empty) subsets of A such that the following hold for each $i \in [t+1]$.

- $|A_i|$ is divisible by k,
- $|A_i| \leq |A|/t + k$,
- every vertex in A' belongs to exactly two sets in $\mathcal{A}$ and every vertex in $A \setminus A'$ belongs to exactly one set in $\mathcal{A}$.

Now, for each $i \in [t+1]$, we have

$$\delta(G - A_i) \geq \delta(G) - |A_i| \geq (1 - 1/k + \sigma)n - n/t - k \geq (1 - 1/k + \sigma - 2/t)n \geq (1 - 1/k)n.$$

Since $V(G)\setminus A_i$ contains at most n vertices, and $|V(G)\setminus A_i|$ is divisible by k, the Hajnal-Szemerédi theorem implies that there exists a collection $\mathcal{K}_i$ of copies of K_k in G covering all the vertices in $V(G)\backslash A_i$ exactly once. For each $i \in [t+1]$, let $E_i := \{V(K) : K \in \mathcal{K}_i\}$. Then $\bigcup_{i=1}^{t+1} E_i$ covers every vertex in $V(G)\backslash A$ exactly $t+1$ times, while it covers vertices in $A \setminus A'$ exactly t times and vertices in A' exactly $t-1$ times. Let $H_{\ell+1}$ be the multi-k-graph on vertex set $V(G)$ with

$$E(H_{\ell+1}) := H_\ell \cup \bigcup_{i=1}^{t+1} E_i.$$

Then the above construction with $(\text{H1})^{\ell}_{3.3.19}$ implies $(\text{H1})^{\ell+1}_{3.3.19}$. Also $(\text{H2})^{\ell}_{3.3.19}$ implies that $\Delta(H_{\ell+1}) = \Delta(H_\ell) + (t+1) \leq (t+1)(\ell+1)$, thus $(\text{H2})^{\ell+1}_{3.3.19}$ holds. If $A' \subsetneq A_{\max}$, then every vertex in $A_{\max} \setminus A'$ is covered exactly t times by $\bigcup_{i=1}^{t+1} E_i$. Thus, by (3.3.14), we have

$$\max_{u\in V(G)} \{p_{H_{\ell+1}}(u)\} = \max_{u\in V(G)} \{p_{H_\ell}(u)\} + t \text{ and } \min_{u\in V(G)} \{p_{H_{\ell+1}}(u)\} = \min_{u\in V(G)} \{p_{H_\ell}(u)\} + t + 1.$$

If $A_{\max} \subseteq A'$, then every vertex in $A_{\max}$ is covered exactly $t-1$ times while every vertex in A is covered either $t-1$ times or t times by $\bigcup_{i=1}^{t+1} E_i$. Thus, by (3.3.14), we have

$$\max_{u\in V(G)} \{p_{H_{\ell+1}}(u)\} = \max_{u\in V(G)} \{p_{H_\ell}(u)\} + t - 1 \text{ and } \min_{u\in V(G)} \{p_{H_{\ell+1}}(u)\} \geq \min_{u\in V(G)} \{p_{H_\ell}(u)\} + t.$$

In both cases, we have

$$\max_{u,v\in V(G)} \left\{p_{H_{\ell+1}}(u) - p_{H_{\ell+1}}(v)\right\} \leq \max_{u,v\in V(G)} \{p_{H_\ell}(u) - p_{H_\ell}(v)\} - 1 \overset{(\text{H3})^{\ell}_{3.3.19}}{\leq} m' - \ell - 1.$$

Thus $(\text{H3})^{\ell+1}_{3.3.19}$ holds.

Next assume that $|A| < k$. Then we take two sets B and C in $V(G)$ such that $B \cap C = A$ and $|B| = |C| = k$. Then similarly as before, we can take two collections E_1 and E_2 of sets of size k such that E_1 covers every vertex in $V(G) \setminus B$ exactly once, and E_2 covers every vertex in $V(G) \setminus C$ exactly once while $G[e] \simeq K_k$ for all $e \in E_1 \cup E_2$. Let $H_{\ell+1}$ be the multi-k-graph with $E(H_{\ell+1}) := H_\ell \cup E_1 \cup E_2$. Then, it is easy to see that both $(\text{H1})^{\ell+1}_{3.3.19}$ and $(\text{H2})^{\ell+1}_{3.3.19}$ hold. Also $E_1 \cup E_2$ covers all vertices in $V(G) \setminus A$ exactly once or twice, while it does not cover the vertices in A. Then as before, by using the fact that $\max_{u\in V(G)}\{p_{H_\ell}(u)\} - \min_{u\in V(G)}\{p_{H_\ell}(u)\} \geq 2$, we can show that $(\text{H3})^{\ell+1}_{3.3.19}$ holds.

Hence, this shows that there exists a hypergraph $H_{m'-1}$ which satisfies $(\text{H1})^{m'-1}_{3.3.19}$–$(\text{H3})^{m'-1}_{3.3.19}$. Let $m'' := \max_{v\in V(G)}\{p_{H_{m'-1}}(v)\}$. Then $(\text{H2})^{m'-1}_{3.3.19}$ implies that $m'' \leq (t+1)m$.

Also, by $(H3)^{m'-1}_{3.3.19}$ every vertex $v \in V(G)$ satisfies $p_{H_{m'-1}}(v) \in \{m''-1, m''\}$. Recall that $\delta(G) \geq (1-1/k)n$ and k divides n. Thus the Hajnal-Szemerédi theorem guarantees a collection E of sets of size k which covers every vertex of G exactly once, while $G[e] \simeq K_k$ for all $e \in E$. Thus, by adding all $e \in E$ to $H_{m'-1}$ exactly $(t+1)m - m''$ times, we obtain a multi-k-graph satisfying $(B1)_{3.3.19}$ and $(B2)_{3.3.19}$. □

The following lemma is due to Komlós, Sárközy and Szemerédi [69]. Assertion $(B3)_{3.3.20}$ is not explicitly stated in [69], but follows immediately from the proof given there (see Section 3.1 in [69]). Given embeddings of graphs H_i and H_j into blown-up k-cliques $Q_i \subseteq G$ and $Q_j \subseteq G$, the 'clique walks' guaranteed by Lemma 3.3.20 will allow us to find suitable connections between (the images of) H_i and H_j in G.

Lemma 3.3.20. *Let $r, k \in \mathbb{N} \setminus \{1\}$. Suppose that R is an r-vertex graph with $\delta(R) \geq \left(1-\frac{1}{k}\right)r+1$. Suppose that Q_1, Q_2 are two not necessarily disjoint subsets of $V(R)$ of size k such that $Q_1 = \{x_1, \dots, x_k\}$ and $Q_2 = \{y_1, \dots, y_k\}$ with $R[Q_1] \simeq K_k$ and $R[Q_2] \simeq K_k$. Then there exists a walk $W = (z_1, \dots, z_t)$ in R satisfying the following.*

$(B1)_{3.3.20}$ *$3k \leq t \leq 3k^3$ and $k \mid t$,*

$(B2)_{3.3.20}$ *for all $i, j \in [t]$ with $|i-j| \leq k-1$, we have $z_i z_j \in E(R)$,*

$(B3)_{3.3.20}$ *for each $i \in [k]$, we have $z_i = x_i$ and $z_{t-k+i} = y_i$.*

The following lemma also can be proved using a simple greedy algorithm. We omit the proof.

Lemma 3.3.21. *Let $\Delta, k, t \in \mathbb{N} \setminus \{1\}$. Let H be a graph with $\Delta(H) \leq \Delta$ and let $X \subseteq V(H)$ be a set with $|X| \geq \Delta^k t$. Then there exists a k-independent set $Y \subseteq X$ of H with $|Y| = t$.*

Lemma 3.3.22. *Let $r, k, q, s \in \mathbb{N} \setminus \{1\}$ with $0 < 1/r \ll 1/k, 1/q \leq 1$. Let R be an r-vertex graph with $\delta(R) \geq (1-\frac{1}{k})r$. Let $\mathcal{F}$ be a multi-$(k-1)$-graph on $V(R)$ with $\Delta(\mathcal{F}) \leq q$ and $E(\mathcal{F}) = \{F_1, \dots, F_s\}$ such that $R[F_i] \simeq K_{k-1}$ for all $i \in [s]$. Then there exists a multi-k-graph $\mathcal{F}^*$ on $V(R)$ with $E(\mathcal{F}^*) = \{F_1^*, \dots, F_s^*\}$ and such that*

(B1)$_{3.3.22}$ $\Delta(\mathcal{F}^*) \leq (k+1)q$,

(B2)$_{3.3.22}$ *for all* $i \in [s]$*, we have* $F_i \subseteq F_i^*$ *and* $R[F_i^*] \simeq K_k$.

Proof. Since $\mathcal{F}$ is a multi-$(k-1)$-graph, we have $s \leq \Delta(\mathcal{F})r/(k-1) \leq qr$. We consider an auxiliary bipartite graph Aux with vertex partition $(E(\mathcal{F}), V(R) \times [kq])$ such that F_i is adjacent to $(v, j) \in V(R) \times [kq]$ if $v \in N_R(F_i)$. For any set X of $k-1$ vertices in R, we have $d_R(X) \geq r/k$. Thus, any vertex F_i of the graph Aux has degree at least $kqd_R(F_i) \geq kq\cdot(r/k) \geq s = |E(\mathcal{F})|$. Thus, the graph Aux contains a matching M covering every $F_i \in E(\mathcal{F})$. For each $(F_i, (v, j)) \in M$, let $F_i^* := F_i \cup \{v\}$. Then (B2)$_{3.3.22}$ holds. On the other hand, for any vertex $v \in V(R)$, we have $d_{\mathcal{F}^*}(v) = d_{\mathcal{F}}(v) + |\{j \in [kq] : d_M((v, j)) = 1\}| \leq d_{\mathcal{F}}(v) + kq \leq (k+1)q$. Thus (B1)$_{3.3.22}$ holds too. □

The final tool we will collect implies that a (k, η)-chromatic η-separable bounded degree graph has a small separator S and a $(k+1)$-colouring in which one colour class is small and only consists of vertices far away from S.

Lemma 3.3.23. *Suppose that* $n, t, \Delta, k \in \mathbb{N}$ *and* $\Delta \geq 2$*. Suppose that* H *is an* η*-separable* n*-vertex graph with* $\Delta(H) \leq \Delta$*. If* H *admits a* $(k+1)$*-colouring with colour classes* $W_0, \dots, W_k$ *with* $|W_0| \leq \eta n$*, then there exists a* $\Delta^{t+2}\eta$*-separator* S *of* H *with* $N_H^t(S) \cap W_0 = \emptyset$.

Proof. As H is η-separable, there exists an η-separator S' of H. Consider $S := (S' \cup N_H^{t+1}(W_0)) \setminus N_H^t(W_0)$. It is obvious that such a choice satisfies $N_H^t(S) \cap W_0 = \emptyset$. Furthermore, as $|W_0| \leq \eta n$ and $\Delta \geq 2$, we have $|S| \leq \Delta^{t+2}\eta n$. Moreover, any component of $H - S$ is either a subset of a a component of $H - S'$ or a subset of $N_H^t(W_0)$. Hence, it has size at most $\Delta^{t+2}\eta n$, and S is a separator as desired. □

3.4 Constructing an appropriate partition of a separable graph

In Section 3.6 we will decompose the host graph G into graphs G_t, F_t and F'_t with $t \in [T]$ for some bounded T. We will also construct an exceptional set V_0 and reservoir sets Res_t. We now need to partition each graph $H \in \mathcal{H}$ so that this partition reflects the above decomposition of G. This will enable us to apply the blow-up lemma for approximate decompositions (Theorem 3.3.15) in Section 3.5. The next lemma ensures that we can prepare each graph $H \in \mathcal{H}$ in an appropriate manner. It gives a partition of $V(H)$ into X, Y, Z, A. Later we will aim to embed the vertices in A into V_0, and vertices in $Y \cup Z$ will be embedded into Res_t using Lemma 3.3.6. Most of the vertices in X will be embedded into a super-regular blown-up K_k-factor in G_t via Theorem 3.3.15, while the remaining vertices of X will be embedded into Res_t. The set Z will contain a suitable separator H_0 of H. The neighbourhoods of the exceptional vertices $a_\ell \in A$ will be allocated to Y. Moreover, (A2)$_{3.4.1}$ and (A3)$_{3.4.1}$ ensure that we allocate them to sets corresponding to (evenly distributed) cliques of R—the latter enables us to satisfy the second part of (B3)$_{3.4.1}$.

Lemma 3.4.1. *Suppose $n, m, r, k, h, \Delta \in \mathbb{N}$ with $0 < 1/n \ll \eta \ll \varepsilon \ll 1/h \ll 1/k, \sigma, 1/\Delta < 1$ and $0 < \eta \ll 1/r < 1$ such that $k \mid r$. Let H be an n-vertex (k, η)-chromatic graph with $e(H) = m$ and $\Delta(H) \leq \Delta$. Let R and Q be graphs with $V(R) = V(Q) = [r]$ such that Q is a union of r/k vertex-disjoint copies of K_k. For some $n' \in [\varepsilon n]$, let $C_1, \dots, C_{n'}$ be subsets of $[r]$ of size $k-1$, and $C^*_1, \dots, C^*_{n'}$ be subsets of $[r]$ of size k. Let $\mathcal{F}$ and $\mathcal{F}^*$ be multi-hypergraphs on $[r]$ with $E(\mathcal{F}) = \{C_1, \dots, C_{n'}\}$ and $E(\mathcal{F}^*) = \{C^*_1, \dots, C^*_{n'}\}$. Suppose that $n_1, \dots, n_r$ are integers. Suppose the following hold.*

(A1)$_{3.4.1}$ $\delta(R) \geq (1 - \frac{1}{k} + \sigma)r$,

(A2)$_{3.4.1}$ *for each $\ell \in [n']$, we have $C_\ell \subseteq C^*_\ell$ and $R[C^*_\ell] \simeq K_k$,*

(A3)$_{3.4.1}$ $\Delta(\mathcal{F}^*) \leq \varepsilon^{2/3} n/r$, (A4)$_{3.4.1}$ *for each* $i \in [r]$*, we have* $n_i = (1 \pm \varepsilon^{1/2})n/r$*, and* $n' + \sum_{i\in[r]} n_i = n$*.*

Then there exists a randomised algorithm which always returns an ordered partition $(X_1, \dots, X_r, Y_1, \dots, Y_r, Z_1, \dots, Z_r, A)$ *of* $V(H)$ *such that* $A = \{a_1, \dots, a_{n'}\}$ *is a* 3*-independent set of* H *and the following hold, where* $X := \bigcup_{i\in[r]} X_i, Y := \bigcup_{i\in[r]} Y_i$*, and* $Z := \bigcup_{i\in[r]} Z_i$*.*

(B1)$_{3.4.1}$ *For each* $\ell \in [n']$*, we have* $d_H(a_\ell) \leq \frac{2(1+1/h)m}{n}$*,*

(B2)$_{3.4.1}$ *for each* $\ell \in [n']$*, we have* $N_H(a_\ell) \subseteq \bigcup_{i\in C_\ell} Y_i \setminus N^1_H(Z)$*,*

(B3)$_{3.4.1}$ $H[X]$ *admits the vertex partition* $(Q, X_1, \dots, X_r)$*, and* $H \setminus E(H[X])$ *admits the vertex partition* $(R, X_1 \cup Y_1 \cup Z_1, \dots, X_r \cup Y_r \cup Z_r)$*,*

(B4)$_{3.4.1}$ *for each* $ij \in E(Q)$*, we have* $e_H(X_i, X_j) = \frac{2m\pm\varepsilon^{1/5}n}{(k-1)r}$*,*

(B5)$_{3.4.1}$ *for each* $i \in [r]$*, we have* $|X_i| + |Y_i| + |Z_i| = n_i \pm \eta^{1/4}n$ *and* $|Y_i| \leq 2\varepsilon^{1/3}n/r$*,*

(B6)$_{3.4.1}$ $N^1_H(X) \setminus X \subseteq Z$ *and* $|Z| \leq 4\Delta^{3k^3}\eta^{0.9}n$*.*

Moreover, the algorithm has the following additional property, where the expectation is with respect to all possible outputs.

(B7)$_{3.4.1}$ *For all* $\ell \in [n']$ *and* $i \in C_\ell$*, we have* $\mathbb{E}[N_H(a_\ell) \cap Y_i] \leq \frac{2(1+1/h)m}{(k-1)n}$*.*

(B1)$_{3.4.1}$ and (B7)$_{3.4.1}$ ensure that each embedding of some H in G does not use too many edges incident to the exceptional set V_0.

Proof. Write $r' := r/k$ and $Q = \bigcup_{s=1}^{r'} Q_s$, where each Q_s is a copy of K_k, and let $\binom{R}{K_k} = \{Q'_1, \dots, Q'_q\}$ be the collection of all copies of K_k in R. By permuting indices if necessary, we may assume that $V(Q'_1) = \{1, \dots, k\}$. Note that $q \leq r^k$. As Q is a K_k-factor on $[r]$, for each $i \in [r]$, there exists a unique $j \in [r']$ such that $i \in Q_j$. For all $s \in [r']$, $s' \in [q]$ and $k' \in [k]$, we define $q_s(k'), q'_{s'}(k') \in [r]$ to be the k'-th smallest number in $V(Q_s)$ and $V(Q'_{s'})$ respectively. Thus

$$V(Q_s) = \{q_s(1), \dots, q_s(k)\} \text{ and } V(Q'_{s'}) = \{q'_{s'}(1), \dots, q'_{s'}(k)\}.$$

For all $s \in [q]$ and $k' \in [k]$, let

$$Q'_{s,k'} := Q'_s \setminus \{q'_s(k')\} \quad \text{and} \quad d_{s,k'} := |\{\ell \in [n'] : C^*_\ell = V(Q'_s) \text{ and } C_\ell = V(Q'_{s,k'})\}|. \quad (3.4.1)$$

Note that for each $i \in [r]$ we have

$$\sum_{s \in [q]: i \in V(Q'_s)} \sum_{k' \in [k]} d_{s,k'} = d_{\mathcal{F}^*}(i) \quad \text{and} \quad \sum_{(s,k') \in [q] \times [k]} d_{s,k'} = n'. \quad (3.4.2)$$

Our strategy is as follows. Consider a $(k+1)$-colouring $(W_0, \dots, W_k)$ of H with $|W_0| \leq \eta n$ and an $\Delta^{3k^3+3}\eta n$-separator S of H guaranteed by Lemma 3.3.23 (applied with $t = 3k^3 + 1$). Thus we can partition the k-chromatic graph $H \setminus W_0$ into $H_0, \dots, H_t$ such that each $H_{t'}$ is small, there are no edges between $H_{t'}$ and $H_{t''}$ whenever $0 \notin \{t', t''\}$ and $V(H_0) = S$. We will distribute the vertices of each graph $H_{t'}$ into $\bigcup_{i \in V(Q_s)} X_i$ or $\bigcup_{i \in V(Q'_s)} (Y_i \cup Z_i)$ for an appropriate s. In particular, $V(H_0)$ will be allocated to $\bigcup_{i \in V(Q'_1)} Z_i = \bigcup_{i \in [k]} Z_i$. As Q'_s and Q_s are copies of K_k in R and Q, respectively, and as $H_{t'}$ is k-chromatic, this would allow us to achieve (B3)$_{3.4.1}$ if we ignore the edges incident to $V(H_0) \cup W_0$. In Steps 5 and 6 we will use 'clique walks' obtained from Lemma 3.3.20 to connect up the $H_{t'}$ with H_0 in a way which respects the colour classes of $H \setminus W_0$. We can thus allocate the vertices in $N_H^{3k^3}(V(H_0))$ in a way that will satisfy (B3)$_{3.4.1}$. Finally, we will allocate the vertices in W_0. As W_0 is far from $V(H_0)$, each vertex in W_0 only has its neighbours in a single $H_{t'}$, hence it will be simple to assign each vertex in W_0 to some Z_i with $i \in [r]$ according to where the vertices of $H_{t'}$ are assigned.

Step 1. Separating H. As H is (k, η)-chromatic, applying Lemma 3.3.23 with $t = 3k^3 + 1$ implies that there exists a partition $(W_0, W_1, \dots, W_k)$ of $V(H)$ into independent sets and an $\eta^{0.9}$-separator S such that

$$|S|, |W_0| \leq \eta^{0.9} n \text{ and } W_0 \cap N_H^{3k^3+1}(S) = \emptyset. \quad (3.4.3)$$

Since S is an $\eta^{0.9}$-separator of H, it follows that there exists a partition $\widetilde{V}_0, \dots, \widetilde{V}_t$ of $V(H)$

with $\widetilde{V}_0 = S$, such that the following hold, where $V_{t'} := \widetilde{V}_{t'} \setminus W_0$ and $H_{t'} := H[V_{t'}]$ for each $t' \in [t] \cup \{0\}$.

(H1)$_{3.4.1}$ $\eta^{-0.9}/2 \leq t \leq 2\eta^{-0.9}$,

(H2)$_{3.4.1}$ $\eta^{0.9}n/2 \leq |V_{t'}| \leq 2\eta^{0.9}n$ for $t' \in [t]$,

(H3)$_{3.4.1}$ for $t' \neq t'' \in [t]$, we have that $E_H(\widetilde{V}_{t'}, \widetilde{V}_{t''}) = \emptyset$, and $m - 2\Delta\eta^{0.9}n \leq \sum_{t' \in [t]} e(H_{t'}) \leq m$.

Indeed, as S is an $\eta^{0.9}$-separator of H, $H \setminus S$ only consists of components of size at most $\eta^{0.9}n$. By letting $\tilde{V}_0 := S$ (and thus $V_0 = S$) and letting each of $\widetilde{V}_1, \dots, \widetilde{V}_t$ be appropriate unions of components of $H \setminus S$, we can ensure that both (H1)$_{3.4.1}$ and (H2)$_{3.4.1}$ hold. By the construction, the first part of (H3)$_{3.4.1}$ holds too. Since there are at most $\Delta(H)|S \cup W_0| \leq 2\Delta\eta^{0.9}n$ edges which are incident to some vertex in $W_0 \cup V_0$, the second part of (H3)$_{3.4.1}$ holds as well.

For each $t' \in [t] \cup \{0\}$ and $k' \in [k]$, we let

$$W_{k'}^{t'} := V_{t'} \cap W_{k'}.$$

Step 2. Choosing the exceptional set A. Let

$$L := \left\{ x \in V(H) : d_H(x) \leq \frac{2(1+1/h)m}{n} \right\}.$$

L contains the 'low degree' vertices within which we will choose A in order to satisfy (B1)$_{3.4.1}$. Note that $2m = \sum_{x \in V(H)} d_H(x) \geq \frac{2(1+1/h)m}{n}(n - |L|)$, thus

$$|L| \geq n/(2h). \tag{3.4.4}$$

For each $t' \in [t]$, let $k(t') \in [k]$ be an index such that

$$|L \cap W_{k(t')}^{t'}| \geq \frac{1}{k}|L \cap V(H_{t'})|. \tag{3.4.5}$$

Such a number $k(t')$ exists as $W_1^{t'}, \dots, W_k^{t'}$ forms a partition of $V_{t'} = V(H_{t'})$.

Now, we choose a partition $\mathcal{H}, \mathcal{H}'_{1,1}, \dots, \mathcal{H}'_{1,k}, \mathcal{H}'_{2,1}, \dots, \mathcal{H}'_{q,k}$ of $\{H_1, \dots, H_t\}$ satisfying the following for each $(s, k') \in [q] \times [k]$.

(H4)$_{3.4.1}$ $v(\mathcal{H}'_{s,k'}) = \varepsilon^{-1/10} d_{s,k'} + 2k\eta^{2/5} n \pm \eta^{2/5} n$ and

$\sum_{t' : H_{t'} \in \mathcal{H}'_{s,k'}} |V(H_{t'}) \cap L| \geq \varepsilon^{-1/11} d_{s,k'} + \eta^{1/2} n$.

We will choose A within the vertex sets of the graphs in $\mathcal{H}'_{1,1}, \ldots, \mathcal{H}'_{q,k}$. Moreover, we will allocate all the other vertices of the graphs in each $\mathcal{H}'_{s,k'}$ to $Y \cup Z$.

Claim 3.4.1. *There exists a partition $\mathcal{H}, \mathcal{H}'_{1,1}, \ldots, \mathcal{H}'_{1,k}, \mathcal{H}'_{2,1}, \ldots, \mathcal{H}'_{q,k}$ of $\{H_1, \ldots, H_t\}$ satisfying (H4)$_{3.4.1}$.*

Proof. For each $t' \in [t]$, we choose $i_{t'}$ independently at random from $[q] \times [k] \cup \{(0,0)\}$ such that for each $(s, k') \in [q] \times [k]$ we have

$$\mathbb{P}[i_{t'} = (s, k')] = \frac{\varepsilon^{-1/10} d_{s,k'}}{n} + 2k\eta^{2/5} \quad \text{and} \quad \mathbb{P}[i_{t'} = (0,0)] = 1 - \frac{\varepsilon^{-1/10} n'}{n} - 2qk^2\eta^{2/5}.$$

An easy calculation based on (3.4.2) shows that this defines a probability distribution. For each $(s, k') \in [q] \times [k]$, we let

$$\mathcal{H} := \{H_{t'} : t' \in [t], i_{t'} = (0,0)\} \quad \text{and} \quad \mathcal{H}'_{s,k'} := \{H_{t'} : t' \in [t], i_{t'} = (s, k')\}.$$

Then it is easy to combine a Chernoff bound (Lemma 3.3.1) with (H1)$_{3.4.1}$, (H2)$_{3.4.1}$, (3.4.4) and the fact that $|V(H)| = n$ to check that the resulting partition satisfies (H4)$_{3.4.1}$ with positive probability. This proves the claim. □

By permuting indices on $[t]$, we may assume that for some $t_* \in [t]$, we have

$$\mathcal{H} = \{H_1, \ldots, H_{t_*}\} \quad \text{and} \quad \bigcup_{(s,k') \in [q] \times [k]} \mathcal{H}'_{s,k'} = \{H_{t_*+1}, \ldots, H_t\}.$$

For each $(s, k') \in [q] \times [k]$, let

$$L_{s,k'} := \bigcup_{t' : H_{t'} \in \mathcal{H}'_{s,k'}} (L \cap W^{t'}_{k(t')}) \setminus N_H^{3k^3+2}(V_0 \cup W_0). \tag{3.4.6}$$

Then by (3.4.3) and (3.4.5) we have

$$|L_{s,k'}| \geq \sum_{t' : H_{t'} \in \mathcal{H}'_{s,k'}} \frac{1}{k} |L \cap V(H_{t'})| - 8\Delta^{3k^3+2} \eta^{0.9} n \overset{(\text{H4})_{3.4.1}}{\geq} \varepsilon^{-1/11} d_{s,k'}/k + \eta^{1/2} n/(2k) \geq \Delta^3 d_{s,k'}.$$

For each $(s,k') \in [q] \times [k]$, we apply Lemma 3.3.21 to $L_{s,k'}$ to obtain a subset of $L_{s,k'}$ with size exactly $d_{s,k'}$ which is 3-independent in H. Write this 3-independent set as

$$\{a_\ell : \ell \in [n'], C^*_\ell = V(Q'_s) \text{ and } C_\ell = V(Q'_{s,k'})\}. \tag{3.4.7}$$

This is possible by (3.4.1) and (3.4.2) and defines vertices $a_1, \dots, a_{n'}$. Let $A := \{a_1, \dots, a_{n'}\}$. By (3.4.6) and $(\mathrm{H3})_{3.4.1}$, A is still a 3-independent set in H. As $a_\ell \in L$, we know that

$$d_H(a_\ell) \leq 2(1 + 1/h)m/n. \tag{3.4.8}$$

Moreover, for $\ell \in [n']$ and $t' \in [t]$, we have the following.

$$\textit{If } a_\ell \in V_{t'}, \textit{ then } t' \in [t] \setminus [t_*] \textit{ and } a_\ell \in W^{t'}_{k(t')} \setminus N^{3k^3+2}_H(V_0 \cup W_0). \tag{3.4.9}$$

In particular, we have $N_H(a_\ell) \cap N^{3k^3+1}_H(V_0 \cup W_0) = \emptyset$. Thus if $a_\ell \in V_{t'}$, then

$$N_H(a_\ell) \subseteq \bigcup_{k'' \in [k] \setminus \{k(t')\}} W^{t'}_{k''} \setminus N^{3k^3+1}_H(V_0 \cup W_0). \tag{3.4.10}$$

Step 3. Allocating the neighbourhood of A. We will allocate $N_H(A)$ to Y. We will achieve this by suitably allocating $V(\mathcal{H}'_{s,k'})$ for each $(s,k') \in [q] \times [k]$. This will allocate $N_H(A)$ via (3.4.10). Note that all choices until now are deterministic. Next we run the following random procedure.

> *For each $t' \in [t] \setminus [t_*]$, let $(s,k') \in [q] \times [k]$ be such that $H_{t'} \in \mathcal{H}'_{s,k'}$, and choose a permutation $\pi_{t'}$ on $[k]$ independently and uniformly at random among all permutations such that $\pi_{t'}(k') = k(t')$.* (3.4.11)

(Note that this is the only place that our choice is random.) Thus one value of $\pi_{t'}$ is fixed, while all other $k-1$ values are chosen at random. We choose $\pi_{t'}$ in this way because we wish to distribute $N_H(a_\ell)$ to $\bigcup_{i \in C_\ell} Y_i$, so that later $(\mathrm{B2})_{3.4.1}$ is satisfied. Setting $\pi_{t'}(k') = k(t')$ will ensure that no vertex in $N_H(a_\ell)$ will be distributed to Y_i with $i \in C^*_\ell \setminus C_\ell$. Moreover, as $\pi_{t'}$ is chosen uniformly at random, $N_H(a_\ell)$ will be distributed to $\bigcup_{i\in C_\ell} Y_i$ in a uniform way, which will guarantee that (B7)$_{3.4.1}$ holds.

Indeed, for $\ell \in [n'], (s,k') \in [q]\times[k]$ and $t' \in [t]\setminus[t_*]$ such that $a_\ell \in L_{s,k'} \cap V_{t'}$, and for any $k'' \in [k]\setminus\{k'\}$, the number $\pi_{t'}(k'')$ is chosen uniformly at random among $[k]\setminus\{k(t')\}$, thus we have

$$\mathbb{E}[|N_H(a_\ell)\cap W^{t'}_{\pi_{t'}(k'')}|] \leq \frac{d_H(a_\ell)}{k-1} \overset{(3.4.8)}{\leq} \frac{2(1+1/h)m}{(k-1)n}. \tag{3.4.12}$$

For each $i\in[r]$, let

$$\tilde{Y}_i := \bigcup_{(s,k'):i=q'_s(k')} \bigcup_{k''\in[k]} \bigcup_{H_{t'}\in\mathcal{H}'_{s,k''}} W^{t'}_{\pi_{t'}(k')}\setminus A \qquad \text{and} \qquad \tilde{Y} := \bigcup_{i\in[r]} \tilde{Y}_i. \tag{3.4.13}$$

Step 4. Allocating the remaining vertices to X and Y. Later the vertices in $\tilde{Y}_i$ will be assigned to Y_i (except those which are too close to V_0 in H, which will be assigned to Z). The sizes of the sets X_i will be almost identical. (Note that because of (B3)$_{3.4.1}$, it is not possible to prescribe different sizes for X_i and X_j if i and j lie in the same copy of K_k in Q.) Thus, in order to ensure (B5)$_{3.4.1}$, we need to decide how many more vertices other than $\tilde{Y}_i$ we will assign to the set Y_i. As part of this we now decide which of the $H_{t'}\in\mathcal{H}$ are allocated to X and which are allocated to Y (again, vertices close to V_0 will be assigned to Z). Note that we have

$$\begin{aligned} |\tilde{Y}_i| &\leq \sum_{(s,k'):i=q'_s(k')} \sum_{k''\in[k]} \sum_{H_{t'}\in\mathcal{H}'_{s,k''}} |H_{t'}| \overset{(H4)_{3.4.1}}{\leq} \sum_{s:i\in V(Q'_s)} \sum_{k''\in[k]} (\varepsilon^{-1/10} d_{s,k''} + 3k\eta^{2/5}n) \\ &\overset{(3.4.2)}{\leq} \varepsilon^{-1/10} d_{\mathcal{F}^*}(i) + 3k^2q\eta^{2/5}n \overset{(A3)_{3.4.1}}{\leq} \varepsilon^{1/2}n/r. \end{aligned} \tag{3.4.14}$$

For each $i\in[r]$, let $\tilde{n} := (1-2\varepsilon^{1/2})n/r$, and

$$\tilde{n}_i := n_i - \tilde{n} - |\tilde{Y}_i| \overset{(A4)_{3.4.1}}{\leq} \frac{\varepsilon^{1/3}n}{(h+1)r}, \text{ then } \tilde{n}_i \overset{(A4)_{3.4.1}}{\geq} \varepsilon^{1/2}n/r - |\tilde{Y}_i| \overset{(3.4.14)}{\geq} 0. \tag{3.4.15}$$

By applying Lemma 3.3.19 with $R, h, \sigma, \varepsilon^{1/3}n/((h+1)r)$ and $\tilde{n}_i$ playing the roles of G, t, σ, m and d_v, respectively, we obtain a multi-k-graph $\mathcal{F}^\#$ on $[r]$ such that for each $Q\in E(\mathcal{F}^\#)$, we have $R[Q] \simeq K_k$, and

for each $i \in [r]$, we have $d_{\mathcal{F}^\#}(i) = \tilde{n}_i + \frac{\varepsilon^{1/3}n}{r} \pm 1$.

$$\tag{3.4.16}$$

This implies

$$\begin{aligned} N \quad &:= \quad \sum_{i\in[r]}(\tilde{n} - \frac{\varepsilon^{1/3}n}{r} + d_{\mathcal{F}^\#}(i)) - |V_0 \cup W_0| \overset{(3.4.15)}{=} \sum_{i\in[r]}(n_i - |\tilde{Y}_i| \pm 1) - |V_0 \cup W_0| \\ &\overset{(\text{A4})_{3.4.1}}{=} \quad n - n' - |\tilde{Y}| - |V_0 \cup W_0| \pm r. \end{aligned} \tag{3.4.17}$$

Note that we have

$$v(\mathcal{H}) = |V(H) \setminus (\tilde{Y} \cup A \cup V_0 \cup W_0)| = N \pm r. \tag{3.4.18}$$

Our target is to assign roughly $d_{\mathcal{F}^\#}(i)$ extra vertices to Y_i in addition to $\tilde{Y}_i$, and assign roughly $\tilde{n} - \frac{\varepsilon^{1/3}n}{r}$ vertices to X_i, and a negligible amount of vertices to Z_i. Then $|X_i| + |Y_i| + |Z_i|$ will be close to n_i as required in (B5)$_{3.4.1}$.

To achieve this, we partition $\mathcal{H} = \{H_1, \dots, H_{t_*}\}$ into $\mathcal{H}_1, \dots, \mathcal{H}_{r'}, \mathcal{H}_1^\#, \dots, \mathcal{H}_q^\#$ satisfying the following for all $i \in [r']$ and $s \in [q]$.

(H5)$_{3.4.1}$ $v(\mathcal{H}_i) = k\tilde{n} - \dfrac{k\varepsilon^{1/3}n}{r} \pm \eta^{2/5}n$ and $e(\mathcal{H}_i) = \dfrac{k(m \pm \varepsilon^{2/7}n)}{r}$,

(H6)$_{3.4.1}$ $v(\mathcal{H}_s^\#) = k \cdot \mathrm{mult}_{\mathcal{F}^\#}(V(Q'_s)) \pm \eta^{2/5}n$.

(Recall that $\mathrm{mult}_{\mathcal{F}^\#}(V(Q'_s))$ denotes the multiplicity of the edge $V(Q'_s)$ in $\mathcal{F}^\#$.) Indeed, such a partition exists by the following claim.

Claim 3.4.2. *There exists a partition $\mathcal{H}_1, \dots, \mathcal{H}_{r'}, \mathcal{H}_1^\#, \dots, \mathcal{H}_q^\#$ of $\{H_1, \dots, H_{t_*}\}$ satisfying* (H5)$_{3.4.1}$– (H6)$_{3.4.1}$.

Proof. For each $t' \in [t_*]$, we choose $i_{t'}$ independently at random from $\{(0,1), \dots, (0,r'), (1,1), \dots, (1,q)\}$ such that for each $i \in [r']$ and $s \in [q]$:

$$\mathbb{P}[i_{t'} = (0,i)] = \frac{k\tilde{n} - \frac{k\varepsilon^{1/3}n}{r} - \frac{k|V_0 \cup W_0|}{r}}{N} \quad \text{and} \quad \mathbb{P}[i_{t'} = (1,s)] = \frac{k \cdot \mathrm{mult}_{\mathcal{F}^\#}(V(Q'_s))}{N}.$$

Since $\sum_{s\in[q]} k \cdot \mathrm{mult}_{\mathcal{F}^\#}(V(Q'_s)) = k|E(\mathcal{F}^\#)| = \sum_{i\in[r]} d_{\mathcal{F}^\#}(i)$, an easy calculation based on (3.4.17) shows that this defines a probability distribution. For all $i \in [r']$ and $s \in [q]$, we let

$$\mathcal{H}_i := \{H_{t'} : t' \in [t_*], i_{t'} = (0,i)\} \quad \text{and} \quad \mathcal{H}^\#_s := \{H_{t'} : t' \in [t_*], i_{t'} = (1,s)\}.$$

Then it is easy to combine a Chernoff bound (Lemma 3.3.1) with (H1)$_{3.4.1}$, (H2)$_{3.4.1}$ and (3.4.18) to check that the resulting partition satisfies (H5)$_{3.4.1}$ and (H6)$_{3.4.1}$ with positive probability. This proves the claim. □

By permuting indices on $[t_*]$, we may assume that for some $t^* \in [t_*]$ we have

$$\bigcup_{i\in[r']} \mathcal{H}_i = \{H_1, \dots, H_{t^*}\} \text{ and } \bigcup_{s\in[q]} \mathcal{H}^\#_s = \{H_{t^*+1}, \dots, H_{t_*}\}.$$

In order to obtain (B3)$_{3.4.1}$–(B5)$_{3.4.1}$, we need to distribute vertices of the graphs in $\mathcal{H}_i$ into $\{X_j : j \in V(Q_i)\}$ and vertices of the graphs in $\mathcal{H}^\#_s$ into $\{Y_j : j \in V(Q'_s)\}$ so that the resulting vertex sets and edge sets are evenly balanced. For this, we define a permutation $\pi_{t'}$ on $[k]$ for each $t' \in [t_*]$ which will determine how we will distribute these vertices. We will choose these permutations $\pi_1, \dots, \pi_{t_*}$ such that the following hold for all $i \in [r']$, $s \in [q]$ and $k' \neq k'' \in [k]$.

(H7)$_{3.4.1}$ $\displaystyle\sum_{t':H_{t'}\in\mathcal{H}_i} |W^{t'}_{\pi_{t'}(k')}| = \tilde{n} - \frac{\varepsilon^{1/3}n}{r} \pm \eta^{2/5}n$ and $\displaystyle\sum_{t':H_{t'}\in\mathcal{H}_i} |E_H(W^{t'}_{\pi_{t'}(k')}, W^{t'}_{\pi_{t'}(k'')})| = \frac{2m \pm \varepsilon^{1/4}n}{(k-1)r}$,

(H8)$_{3.4.1}$ $\displaystyle\sum_{t':H_{t'}\in\mathcal{H}^\#_s} |W^{t'}_{\pi_{t'}(k')}| = \mathrm{mult}_{\mathcal{F}^\#}(V(Q'_s)) \pm \eta^{2/5}n.$

To see that such permutations exist we consider for each $t' \in [t_*]$ a permutation $\pi_{t'} : [k] \to [k]$ chosen independently and uniformly at random. Then, by a Chernoff bound (Lemma 3.3.1) combined with (H1)$_{3.4.1}$ and (H2)$_{3.4.1}$, it is easy to check that $\pi_1, \dots, \pi_{t_*}$ satisfy (H7)$_{3.4.1}$ and (H8)$_{3.4.1}$ with positive probability.

Step 5. Clique walks. Recall that V_0 is a separator of both H and $H \setminus W_0$. The vertices in V_0 will be allocated to the sets $Z_1, \dots, Z_k$ which initially correspond to the clique $Q'_1 \subseteq R$

(recall that $V(Q'_1) = \{1, \dots, k\}$). We now identify an underlying structure in R that will be used in Step 6 to ensure that while allocating $V(H) \setminus (V_0 \cup W_0 \cup A)$ to X, Y and Z, we do not violate the vertex partition admitted by R (c.f. (B3)$_{3.4.1}$). (This is a particular issue when considering edges between separator vertices and the rest of the partition.)

To illustrate this, let $s \in S$ be a separator vertex allocated to $Z_{k'}$. Let x be some vertex in some $H_{t'}$ with $xs \in E(H)$. Suppose $H_{t'}$ is assigned to some clique $Q_i \subseteq Q$ and that this would assign x to some set $X_{i'}$, where $i' \in V(Q_i)$. Furthermore, suppose $i'k'$ is not an edge in R. We cannot simply reassign x to another set X_j to obey the vertex partition admitted by R without also considering the neighbourhood of x in $H_{t'}$. To resolve this, we apply Lemma 3.3.20 to obtain a suitable 'clique walk' P between Q'_1 and Q_i, i.e. the initial sement of P is $V(Q'_1)$, its final segment is $V(Q_i)$ and each segment of k consecutive vertices in P corresponds to a k-clique in R. We initially assign x to a set $Z_{k''}$ for some $k'' \in [k] \setminus \{k'\}$. We then assign the vertices which are close to x to some $Z_{k'''}$, where the choice of $k''' \in [r]$ is determined by P. (In order to connect Y to V_0, we also choose similar clique walks starting with Q'_1 and ending with Q'_s for each $s \in [q]$.)

To define the clique walks formally, for each $t' \in [t]$, let

$$P_{t'} := \begin{cases} Q_i & \text{if } H_{t'} \in \mathcal{H}_i \text{ for some } i \in [r'], \\ Q'_s & \text{if } H_{t'} \in \mathcal{H}^{\#}_s \text{ for some } s \in [q], \\ Q'_s & \text{if } H_{t'} \in \mathcal{H}'_{s,k'} \text{ for some } (s,k') \in [q] \times [k], \end{cases} \quad \text{and} \quad \begin{array}{l} \{p_{t'}(1), \dots, p_{t'}(k)\} := P_{t'}, \\ \text{where } p_{t'}(1) < \dots < p_{t'}(k). \end{array} \tag{3.4.19}$$

By using (A1)$_{3.4.1}$, we can apply Lemma 3.3.20 for each $t' \in [t]$ with $V(Q'_1)$ and $V(P_{t'})$ playing the roles of Q_1 and Q_2 in order to obtain a walk $j(t',1), \dots, j(t', b_{t'}k)$ in R such that

for all distinct $i, i' \in [b_{t'}k]$ with $|i - i'| \le k-1$, we have $j(t',i)j(t',i') \in E(R)$, and for each $k' \in [k]$ we have $j(t',k') = \pi_{t'}(k')$ and $j(t',(b_{t'}-1)k+k') = p_{t'}(k')$. (3.4.20)

Moreover, for each $t' \in [t]$, we have

$$3 \le b_{t'} \le 3k^2. \tag{3.4.21}$$

As described above we will later distribute some vertices of $V_{t'} \cap N^{(b_{t'}-1)k}(V_0)$ to $\bigcup_{k' \in [(b_{t'}-1)k]} Z_{j(t',k')}$ so that we can ensure (B3)3.4.1 and (B6)3.4.1 hold.

Step 6. Iterative construction of the partition. Now, we will distribute the vertices of each $H_{t'}$ into $X_1, \dots, X_r, Y_1, \dots, Y_r, Z_1, \dots, Z_r$ in such a way that (B1)3.4.1–(B7)3.4.1 hold. (In particular, as discussed earlier, we will have $\widetilde{Y}_i \subseteq Y_i$.) To achieve this, for each $t' = 0, 1, \dots, t$, we iteratively define sets $X_1^{t'}, \dots, X_r^{t'}$, $Y_1^{t'}, \dots, Y_r^{t'}$, $Z_1^{t'}, \dots, Z_r^{t'}$. First, for each $k' \in [k]$, let $Z_{k'}^0 := W_{k'}^0$ and for all $i \in [r]$ and $i' \in [r] \setminus [k]$, let

$$X_i^0 := \emptyset, \;\; Y_i^0 := \emptyset \;\text{ and }\; Z_{i'}^0 := \emptyset.$$

We will write

$$V^{t'} := \bigcup_{t''=0}^{t'} V_{t''}, \quad X^{t'} := \bigcup_{i \in [r]} X_i^{t'}, \quad Y^{t'} := \bigcup_{i \in [r]} Y_i^{t'} \quad \text{and} \quad Z^{t'} := \bigcup_{i \in [r]} Z_i^{t'}.$$

Assume that for some $t' \in [t]$, we have already defined a partition $X_1^{t'-1}, \dots, X_r^{t'-1}$, $Y_1^{t'-1}, \dots, Y_r^{t'-1}$, $Z_1^{t'-1}, \dots, Z_r^{t'-1}$ of $V^{t'-1}$ satisfying the following.

(Z1)$^{t'-1}_{3.4.1}$ For all $i' \in [r']$ and $i \in V(Q_{i'})$, let k' be so that $i = q_{i'}(k')$. Then we have (where $b_{t''}$ below is the length of the walk defined in (3.4.20))

$$\bigcup_{t'' \in [t'-1]: H_{t''} \in \mathcal{H}_{i'}} W^{t''}_{\pi_{t''}(k')} \setminus N_H^{(b_{t''}-1)k}(V_0) \subseteq X_i^{t'-1} \subseteq \bigcup_{t'' \in [t'-1]: H_{t''} \in \mathcal{H}_{i'}} W^{t''}_{\pi_{t''}(k')},$$

(Z2)$^{t'-1}_{3.4.1}$ for each $i \in [r]$, we have

$$\bigcup_{k' \in [k]} \bigcup_{\substack{t'' \in [t'-1] \setminus [t^*]: \\ p_{t''}(k') = i}} W^{t''}_{\pi_{t''}(k')} \setminus N_H^{(b_{t''}-1)k}(V_0) \subseteq Y_i^{t'-1} \subseteq \bigcup_{k' \in [k]} \bigcup_{\substack{t'' \in [t'-1] \setminus [t^*]: \\ p_{t''}(k') = i}} W^{t''}_{\pi_{t''}(k')},$$

(Z3)$^{t'-1}_{3.4.1}$ for all $ij \notin E(Q)$, we have $e_H(X_i^{t'-1}, X_j^{t'-1}) = 0$,

(Z4)$^{t'-1}_{3.4.1}$ for all $ij \notin E(R)$, we have $e_H(X_i^{t'-1}, Z_j^{t'-1}) = e_H(Y_i^{t'-1}, Z_j^{t'-1}) = e_H(Y_i^{t'-1}, Y_j^{t'-1}) = e_H(Z_i^{t'-1}, Z_j^{t'-1}) = 0$,

$(\mathrm{Z5})^{t'-1}_{3.4.1}$ $N^1_H(X^{t'-1}) \setminus X^{t'-1} \subseteq Z^{t'-1} \subseteq N^{3k^3}_H(V_0)$,

$(\mathrm{Z6})^{t'-1}_{3.4.1}$ for each $k' \in [k]$, we have $W^0_{k'} \subseteq Z^{t'-1}_{k'}$,

$(\mathrm{Z7})^{t'-1}_{3.4.1}$ for each $t'' \in [t'-1]$, we have $|\{i \in [r] : (X^{t'-1}_i \cup Y^{t'-1}_i) \cap V_{t''} \neq \emptyset\}| \leq k$.

Using that Q'_1 is a copy of K_k in R and $V(Q'_1) = \{1, \ldots, k\}$, it is easy to see that $(\mathrm{Z1})^0_{3.4.1}$–$(\mathrm{Z7})^0_{3.4.1}$ hold with the above definition of X^0_i, Y^0_i, Z^0_i. We now distribute the vertices of $H_{t'}$ by setting

$$X^{t'}_i := \begin{cases} X^{t'-1}_i \cup \left(W^{t'}_{\pi_{t'}(k')} \setminus N^{(b_{t'}-2)k+k'}_H(V_0)\right) & \text{if } t' \in [t^*] \text{ and } i = p_{t'}(k') \text{ for some } k' \in [k], \\ X^{t'-1}_i & \text{otherwise,} \end{cases}$$

$$Y^{t'}_i := \begin{cases} Y^{t'-1}_i \cup \left(W^{t'}_{\pi_{t'}(k')} \setminus N^{(b_{t'}-2)k+k'}_H(V_0)\right) & \text{if } t' \in [t] \setminus [t^*] \text{ and } i = p_{t'}(k') \text{ for some } k' \in [k], \\ Y^{t'-1}_i & \text{otherwise,} \end{cases}$$

$$Z^{t'}_i := Z^{t'-1}_i \cup \bigcup_{\substack{(b,k') \in [b_{t'}-1] \times [k]: \\ i = j(t', (b-1)k+k')}} \left(W^{t'}_{\pi_{t'}(k')} \cap \left(N^{(b-1)k+k'}_H(V_0) \setminus N^{(b-2)k+k'}_H(V_0)\right)\right).$$

Let $H' := H \setminus W_0$. Recall that $N^{3k^3+1}_H(V_0)$ does not contain any vertex in W_0 (see (3.4.3)). Hence $N^i_H(V_0) = N^i_{H'}(V_0)$ for any $i \leq 3k^3 + 1$.

Note that the above definition of $X^{t'}_i, Y^{t'}_i, Z^{t'}_i$ uniquely distributes all vertices of $V^{t'}$. Indeed, first note that either $Y^{t'}_i = Y^{t'-1}_i$ for all $i \in [r]$ or $X^{t'}_i = X^{t'-1}_i$ for all $i \in [r]$ depending on whether $H_{t'} \in \mathcal{H}_c$ for some $c \in [r']$ (in which case $t' \in [t^*]$) or $H_{t'} \in \mathcal{H}^{\#}_s$ for some $s \in [q]$ or $H_{t'} \in \mathcal{H}'_{s,k'}$ for some $(s, k') \in [q] \times [k]$ (in the latter two cases we have $t' \in [t] \setminus [t^*]$). Now, consider $W^{t'}_{k''} \cap (N^a_H(V_0) \setminus N^{a-1}_H(V_0))$ for $k'' \in [k]$ and $a \in \mathbb{N}$. Note $k'' = \pi_{t'}(k')$ for some $k' \in [k]$. Then either $a > (b_{t'} - 2)k + k'$ or $a \in [(b'-1)k + k'] \setminus [(b'-2)k + k']$ for some unique $b' \in [b_{t'} - 1]$. Thus indeed every vertex of $V^{t'}$ belongs to exactly one of $X^{t'}_i$ or $Y^{t'}_i$ or $Z^{t'}_i$.

It is easy to see that the above definition with (3.4.21), $(\mathrm{Z1})^{t'-1}_{3.4.1}$ and $(\mathrm{Z2})^{t'-1}_{3.4.1}$ implies $(\mathrm{Z1})^{t'}_{3.4.1}$ and $(\mathrm{Z2})^{t'}_{3.4.1}$. Also, $(\mathrm{Z7})^{t'}_{3.4.1}$ is obvious from the construction. Moreover, $(\mathrm{Z3})^{t'-1}_{3.4.1}$ and $(\mathrm{H3})_{3.4.1}$ imply $(\mathrm{Z3})^{t'}_{3.4.1}$ while $(\mathrm{Z6})^{t'-1}_{3.4.1}$ implies $(\mathrm{Z6})^{t'}_{3.4.1}$. Similarly, we have $e_H(Y^{t'}_i, Y^{t'}_j) = 0$ if

$ij \notin E(R)$. We now verify the remaining assertions of $(\mathrm{Z4})^{t'}_{3.4.1}$. First suppose that

$$E_H(X_i^{t'}, Z_{i'}^{t'}) \setminus E_H(X_i^{t'-1}, Z_{i'}^{t'-1}) \neq \emptyset \text{ or } E_H(Y_i^{t'}, Z_{i'}^{t'}) \setminus E_H(Y_i^{t'-1}, Z_{i'}^{t'-1}) \neq \emptyset.$$

Then by $(\mathrm{H3})_{3.4.1}$, we have $i = p_{t'}(k')$ for some $k' \in [k]$ and $i' = j(t', (b-1)k + k'')$ for some $k'' \in [k]$ and $b \in [b_{t'} - 1]$, and H contains an edge between

$$W^{t'}_{\pi_{t'}(k')} \setminus N_H^{(b_{t'}-2)k+k'}(V_0) \text{ and } W^{t'}_{\pi_{t'}(k'')} \cap N_H^{(b-1)k+k''}(V_0).$$

This means that $(b_{t'} - 2)k + k' \leq (b-1)k + k''$. Thus $b = b_{t'} - 1$ and $k' \leq k''$. Moreover, since $W^{t'}_{\pi_{t'}(k')}$ is an independent set of H, we have $k' \neq k''$. Since (3.4.20) implies that $i = p_{t'}(k') = j(t', (b_{t'} - 1)k + k')$ and $i' = j(t', (b_{t'} - 2)k + k'')$ with $0 < (b_{t'} - 1)k + k' - ((b_{t'} - 2)k + k'') < k$, again this with (3.4.20) implies that $ii' \in E(R)$. Now suppose that

$$xy \in E_H(Z_i^{t'}, Z_{i'}^{t'}) \setminus E_H(Z_i^{t'-1}, Z_{i'}^{t'-1}) \text{ with } x, y \notin V_0.$$

Then by $(\mathrm{H3})_{3.4.1}$, we have $i = j(t', (b-1)k + k')$ and $i' = j(t', (b'-1)k + k'')$ for some $b, b' \in [b_t - 1]$ and $k' \neq k'' \in [k]$. However, the definition of $Z_i^{t'}$ implies that such an edge only exists when $|((b-1)k + k') - ((b'-1)k + k'')| \leq k - 1$. In this case, (3.4.20) implies that $ii' \in E(R)$. Finally, suppose that

$$xy \in E_H(Z_i^{t'}, Z_{i'}^{t'}) \setminus E_H(Z_i^{t'-1}, Z_{i'}^{t'-1}) \text{ with } x \in V_0 \cap Z_i^{t'}.$$

Then the definition of $Z_i^{t'}$ implies that $i \in [k]$, $x \in W_i^0$ and $i' = j(t', k')$ for some $k' \in [k]$. (3.4.20) implies that $j(t', k') = \pi_{t'}(k')$. As $W^0_{\pi_{t'}(k')} \cup W^{t'}_{\pi_{t'}(k')}$ is an independent set of H, we have $i \neq \pi_{t'}(k')$. However, as $R[[k]] = R[V(Q_1')] \simeq K_k$, we know that $ii' \in E(R)$. Thus $(\mathrm{Z4})^{t'}_{3.4.1}$ holds. By the definition of $X_i^{t'}$ and $Z_i^{t'}$ with (3.4.21), it is obvious that $(\mathrm{Z5})^{t'}_{3.4.1}$ holds too.

Thus, by repeating this, we obtain a partition $X_1^t, \dots, X_r^t, Y_1^t, \dots, Y_r^t, Z_1^t, \dots, Z_r^t$ of $V(H) \setminus W_0$ satisfying $(\mathrm{Z1})^t_{3.4.1}$–$(\mathrm{Z7})^t_{3.4.1}$. For each $i \in [r]$, let

$$X_i := X_i^t, \ \ X := X^t, \ \ Y_i := Y_i^t \setminus A, \ \ Y := Y^t \setminus A, \ \ Z_i' := Z_i^t \text{ and } Z' := Z^t.$$

Note that $A \subseteq Y^t$ by (3.4.9) and $(\mathrm{Z2})^t_{3.4.1}$. Moreover, X, Y, Z', A forms a partition of $V(H) \setminus W_0$. Now we consider the vertices in W_0. For each $w \in W_0$, let

$$I_w := \{i \in [r] : N_H(w) \cap (X_i \cup Y_i) \neq \emptyset\}.$$

By (3.4.3), we have $W_0 \cap V_0 = \emptyset$. Hence, for each vertex $w \in W_0$, there exists $t' \in [t]$ such that $w \in \widetilde{V}_{t'}$. As W_0 is an independent set, (3.4.3) with $(\mathrm{H3})_{3.4.1}$ implies $N_H(w) \subseteq V_{t'}$. This with $(\mathrm{Z7})^t_{3.4.1}$ implies that $|I_w| \leq k$. As $|N_R(I_w)| > 0$ by $(\mathrm{A1})_{3.4.1}$, we can assign w to Z'_i for some $i \in N_R(I_w)$. Let $Z_1, \dots, Z_r, Z$ be the sets obtained from $Z'_1, \dots, Z'_r, Z'$ by assigning all vertices in W_0 in this way. By (3.4.3), (3.4.9) and $(\mathrm{Z5})^t_{3.4.1}$ for each $w \in W_0$ we have $N_H(w) \subseteq X \cup Y$. Thus

$$\text{for all } i \in [r],\ w \in W_0 \cap Z_i \text{ and } x \in N_H(w), \text{ we have } x \in X_j \cup Y_j \text{ for some } j \in N_R(i). \tag{3.4.22}$$

The sets X, Y, Z, A now form a partition of $V(H)$.

Step 7. Checking the properties of the partition. We now verify that this partition satisfies $(\mathrm{B1})_{3.4.1}$-$(\mathrm{B7})_{3.4.1}$. Note that (3.4.8) implies $(\mathrm{B1})_{3.4.1}$. Consider any $\ell \in [n']$, and let $t' \in [t] \setminus [t_*]$ and $(s, k') \in [q] \times [k]$ be such that $a_\ell \in H_{t'} \in \mathcal{H}'_{s,k'}$. Then

$$\begin{aligned} N_H(a_\ell) &\overset{(3.4.10)}{\subseteq} \bigcup_{k'' \in [k] \setminus \{k(t')\}} W^{t'}_{k''} \setminus N_H^{3k^3+1}(V_0 \cup W_0) \overset{(3.4.11)}{=} \bigcup_{k'' \in [k] \setminus \{k'\}} W^{t'}_{\pi_{t'}(k'')} \setminus N_H^{3k^3+1}(V_0 \cup W_0) \\ &\overset{(\mathrm{Z2})^t_{3.4.1}, (\mathrm{Z5})^t_{3.4.1}}{\subseteq} \bigcup_{k'' \in [k] \setminus \{k'\}} Y_{p_{t'}(k'')} \setminus N_H^1(Z) \overset{(3.4.1),(3.4.19)}{=} \bigcup_{i \in V(Q'_{s,k'})} Y_i \setminus N_H^1(Z) \overset{(3.4.7)}{=} \bigcup_{i \in C_\ell} Y_i \setminus N_H^1(Z). \end{aligned}$$

This proves $(\mathrm{B2})_{3.4.1}$. Moreover, whenever ℓ, t' and (s, k') are as in the proof of $(\mathrm{B2})_{3.4.1}$, for each $j' \in C_\ell$, we have $j' = p_{t'}(k'')$ for some $k'' \in [k] \setminus \{k'\}$. Thus by (3.4.10) and $(\mathrm{Z2})^t_{3.4.1}$, we have

$$\mathbb{E}[|N_H(a_\ell) \cap Y_{j'}|] \leq \mathbb{E}[|N_H(a_\ell) \cap W^{t'}_{\pi_{t'}(k'')}|] \overset{(3.4.12)}{\leq} \frac{2(1 + 1/h)m}{(k-1)n}.$$

This proves $(\mathrm{B7})_{3.4.1}$.

Properties $(\mathrm{Z3})^t_{3.4.1}$, $(\mathrm{Z4})^t_{3.4.1}$, $(\mathrm{Z5})^t_{3.4.1}$ and (3.4.22) imply $(\mathrm{B3})_{3.4.1}$.

For each $ij \in E(Q)$, let $s \in [r']$ and $k', k'' \in [k]$ be such that $i = q_s(k')$ and $j = q_s(k'')$. Thus

$$e_H(X_i, X_j) \overset{(\mathrm{H3})_{3.4.1},(\mathrm{Z1})^t_{3.4.1}}{=} \sum_{t' \in [t^*]: H_{t'} \in \mathcal{H}_s} |E_H(W^{t'}_{\pi_{t'}(k')}, W^{t'}_{\pi_{t'}(k'')})| \pm \Delta |N_H^{3k^3}(V_0)| \overset{(\mathrm{H2})_{3.4.1},(\mathrm{H7})_{3.4.1}}{=} \frac{2m \pm \varepsilon^{1/5} n}{(k-1)r}.$$

Thus $(\mathrm{B4})_{3.4.1}$ holds. Moreover, given $i \in [r]$, let $s \in [r']$ and $k' \in [k]$ be such that $i = q_s(k')$. Then

$$|X_i| \overset{(\mathrm{Z1})^t_{3.4.1}}{=} \sum_{t' \in [t^*]: H_{t'} \in \mathcal{H}_s} |W^{t'}_{\pi_{t'}(k')}| \pm |N_H^{3k^3}(V_0)| \overset{(\mathrm{H7})_{3.4.1}}{=} \tilde{n} - \varepsilon^{1/3} n/r \pm \eta^{1/3} n.$$

Similarly, for $i \in [r]$, since by (3.4.9) the vertices of A only belong to $V(H_{t'})$ for $t' \in [t] \setminus [t_*]$,

$$\begin{aligned}
|Y_i| &\overset{(\mathrm{Z2})^t_{3.4.1}}{=} \sum_{(t',k'): p_{t'}(k') = i, t' \in [t] \setminus [t^*]} |W^{t'}_{\pi_{t'}(k')} \setminus A| \pm |N_H^{3k^3}(V_0)| \\
&\overset{(3.4.19)}{=} \sum_{(s,k'): q'_s(k') = i} \sum_{t': H_{t'} \in \mathcal{H}^{\#}_s} |W^{t'}_{\pi_{t'}(k')}| + \sum_{(s,k'): q'_s(k') = i} \sum_{k'' \in [k]} \sum_{t': H_{t'} \in \mathcal{H}'_{s,k''}} |W^{t'}_{\pi_{t'}(k')} \setminus A| \pm \eta^{1/2} n \\
&\overset{(\mathrm{H8})_{3.4.1},(3.4.13)}{=} \sum_{(s,k'): q'_s(k') = i} \mathrm{mult}_{\mathcal{F}^{\#}}(V(Q'_s)) + |\tilde{Y}_i| \pm 2q\eta^{2/5} n = d_{\mathcal{F}^{\#}}(i) + |\tilde{Y}_i| \pm 2q\eta^{2/5} n \\
&\overset{(3.4.15),(3.4.16)}{=} n_i - \tilde{n} + \varepsilon^{1/3} n/r \pm \eta^{1/3} n.
\end{aligned}$$

Together with (3.4.3), $(\mathrm{Z5})^t_{3.4.1}$ and $(\mathrm{H2})_{3.4.1}$, this now implies that for each $i \in [r]$

$$|X_i| + |Y_i| + |Z_i| = n_i \pm \eta^{1/4} n.$$

Also, the definition of $\tilde{n}$ with $(\mathrm{A4})_{3.4.1}$ implies that $|Y_i| \leq 2\varepsilon^{1/3} n/r$. Thus $(\mathrm{B5})_{3.4.1}$ holds. Finally, (3.4.3) and $(\mathrm{Z5})_{3.4.1}$ imply $(\mathrm{B6})_{3.4.1}$. □

3.5 Packing graphs into a super-regular blow-up

In this section, we prove our main lemma. Roughly speaking, this lemma says the following. Suppose we have disjoint vertex sets V, Res_t and V_0 and suppose that we have a super-regular K_k-factor blow-up $G[V]$ on vertex set V, and suitable graphs $G[Res_t]$, $G[V, Res_t]$, $F[V, Res_t]$

and $F'[Res_t, V_0]$ are also provided. Then we can pack an appropriate collection $\mathcal{H}$ of graphs into $G \cup F \cup F'$. Here V_0 is the exceptional set obtained from an application of Szemerédi's regularity lemma and Res_t is a suitable 'reservoir' set where V_0 is much smaller than Res_t, which in turn is much smaller than V. The k-cliques provided by the multi-k-graph $\mathcal{C}_t^*$ below will allow us to find a suitable embedding of the neighbours of the vertices mapped to V_0. When we apply Lemma 3.5.1 in Section 6, the reservoir set Res_t will play the role of the set $U \cup U_0$ below. U_0 will correspond to a set of exceptional vertices in Res_t. (A9)$_{3.5.1}$ will allow us to embed the neighbours of the vertices mapped to U_0.

Note that the packing ϕ is designed to cover most of the edges of the blown-up K_k-factor $G[V]$, but only covers a small proportion of the edges of G incident to U. (A7)$_{3.5.1}$ provides the edges incident to the vertices mapped to V_0, and (A8)$_{3.5.1}$ allows us to embed the neighbourhoods of these vertices.

Lemma 3.5.1. *Suppose $n, n', k, \Delta, r, T \in \mathbb{N}$ with $0 < 1/n, 1/n' \ll \eta \ll \varepsilon \ll 1/T \ll \alpha \ll d \ll 1/k, \sigma, \nu, 1/\Delta < 1$ and $\eta \ll 1/r \ll \sigma$ and $k \mid r$. Suppose that R and Q are graphs with $V(R) = V(Q) = [r]$ such that Q is a union of r/k vertex-disjoint copies of K_k. Suppose that $V_0, \dots, V_r, U_0, \dots, U_r$ is a partition of a set of n vertices such that $|V_0| \le \varepsilon n$, $|U_0| \le \varepsilon n$ and for all $i \in [r]$*

$$n' = |V_i| = \frac{(1 - 1/T \pm 2\varepsilon)n}{r} \quad \text{and} \quad |U_i| = \frac{(1 \pm 2\varepsilon)n}{Tr}.$$

Let $V := \bigcup_{i \in [r]} V_i$ and $U := \bigcup_{i \in [r]} U_i$. Suppose that G, F, F' are edge-disjoint graphs such that $V(G) = V \cup U \cup U_0$, F is a bipartite graph with vertex partition (V, U), and F' is a bipartite graph with vertex partition (V_0, U) such that $F' = \bigcup_{t \in [T]} \bigcup_{v \in V_0} F'_{v,t}$, where all the $F'_{v,t}$ are pairwise edge-disjoint stars with centre v.

Suppose that $\mathcal{H}$ is a collection of (k, η)-chromatic η-separable graphs on n vertices, and for each $t \in [T]$ we have a multi-$(k-1)$-graph $\mathcal{C}_t$ on $[r]$ and a multi-k-graph $\mathcal{C}_t^*$ on $[r]$ with $E(\mathcal{C}_t) = \{C_{v,t} : v \in V_0\}$ and $E(\mathcal{C}_t^*) = \{C_{v,t}^* : v \in V_0\}$. Assume the following hold.

(A1)$_{3.5.1}$ *For each $H \in \mathcal{H}$, we have $\Delta(H) \le \Delta$ and $e(H) \ge n/4$,* (A2)$_{3.5.1}$ $n^{7/4} \leq e(\mathcal{H}) \leq (1-\nu)(k-1)\alpha n^2/(2r)$,

(A3)$_{3.5.1}$ *$G[V]$ is $(T^{-1/2}, \alpha)$-super-regular with respect to the vertex partition $(Q, V_1, \ldots, V_r)$,*

(A4)$_{3.5.1}$ *for each $ij \in E(R)$, the graphs $G[V_i, U_j]$ and $G[U_i, U_j]$ are both $(\varepsilon^{1/50}, (d^3))^+$-regular,*

(A5)$_{3.5.1}$ $\delta(R) \geq (1 - 1/k + \sigma)r$,

(A6)$_{3.5.1}$ *for all $ij \in E(Q)$ and $u \in U_i$, we have $d_{F,V_j}(u) \geq d^3 n'$,*

(A7)$_{3.5.1}$ *for all $v \in V_0$ and $t \in [T]$ and $i \in C_{v,t}$, we have $d_{F'_{v,t},U_i}(v) \geq (1-d)\alpha|U_i|$,*

(A8)$_{3.5.1}$ *for all $v \in V_0$ and $t \in [T]$, we have $C_{v,t} \subseteq C^*_{v,t}$, $R[C^*_{v,t}] \simeq K_k$, and $\Delta(\mathcal{C}^*_t) \leq \frac{\varepsilon^{3/4} n}{r}$,*

(A9)$_{3.5.1}$ *for each $u \in U_0$, we have*

$$|\{i \in [r] : d_{G,V_j}(u) \geq d^3 n' \text{ for all } j \in N_Q(i)\}| > \varepsilon^{1/4} r.$$

Then there exists a packing ϕ of $\mathcal{H}$ into $G \cup F \cup F'$ such that

(B1)$_{3.5.1}$ $\Delta(\phi(\mathcal{H})) \leq 4k\Delta\alpha n/r$,

(B2)$_{3.5.1}$ *for each $u \in U$, we have $d_{\phi(\mathcal{H})\cap G}(u) \leq 2\Delta\varepsilon^{1/8} n/r$,*

(B3)$_{3.5.1}$ *for each $i \in [r]$, we have $e_{\phi(\mathcal{H})\cap G}(V_i, U \cup U_0) < \varepsilon^{1/2} n^2/r^2$.*

Roughly, the proof of Lemma 3.5.1 will proceed as follows. In Step 1 we define a partition of U_0 and an auxiliary digraph D. In Step 2 we define a partition of each $H \in \mathcal{H}$. For each graph $H \in \mathcal{H}$ we apply Lemma 3.4.1 to partition $V(H)$ into X^H, Y^H, Z^H, A^H. We will embed A^H into V_0 and the remainder of H into $V \cup U \cup U_0$. In Step 3, we apply Lemma 3.3.6 to find an appropriate function ϕ' packing $\{H[Y^H \cup Z^H \cup A^H] : H \in \mathcal{H}\}$ into $G[U] \cup F'$. Guided by the auxiliary digraph D, in Step 4 we modify the partition by removing a suitable W^H from X^H (so that we can later embed $X^H \backslash W^H$ into V). We will also find a function ϕ'' packing $\{H[W^H] : H \in \mathcal{H}\}$ into $G[U]$ in an appropriate way, which ensures that later we can also pack $\{H[X^H \setminus W^H, W^H] : H \in \mathcal{H}\}$ into $F[V,U] \cup G[V,U]$. In Step 5 we will partition $\mathcal{H}$ into subcollections $\mathcal{H}_{1,1}, \dots, \mathcal{H}_{T,w}$ and use Lemma 3.3.14 to pack $\{H[X^H \setminus W^H] : H \in \mathcal{H}_{t,w'}\}$ into an internally q-regular graph $H_{t,w'}$ (for some suitable q). Finally, in Step 6 we apply the blow-up lemma for approximate decompositions (Theorem 3.3.15) to pack $\{H_{t,w'} : t \in [T], w' \in [w]\}$ into $G[V]$ such that the packing obtained is consistent with $\phi' \cup \phi''$.

Proof. Let $r' := r/k$ and $Q_1, \dots, Q_{r'}$ be the copies of K_k in Q. Let $n_0 := |V_0|$ and $V_0 =: \{v_1, \dots, v_{n_0}\}$. By $(A1)_{3.5.1}$, for each $H \in \mathcal{H}$, we have

$$e(H) \leq \Delta n. \tag{3.5.1}$$

Moreover,

$$\kappa := |\mathcal{H}| \overset{(A1)_{3.5.1},(A2)_{3.5.1}}{\leq} 2(1-\nu)(k-1)\alpha n/r. \tag{3.5.2}$$

Step 1. Partition of U_0 and the construction of an auxiliary digraph D. In Step 2, we will find a partition of each $H \in \mathcal{H}$ which closely reflects the structure of G. However we need the partitions to match up exactly. The following auxiliary graph will enable us to carry out this adjustment in Step 4. Let D be the directed graph with $V(D) = [r]$ and

$$E(D) = \{\vec{ij} : i \neq j \in [r], N_Q(i) \subseteq N_R(j)\}. \tag{3.5.3}$$

For each $ij \in E(R)$, we let

$$U_i(j) := \{u \in U_i : d_{G,V_j}(u) \geq (d^3 - \varepsilon^{1/50})n'\}.$$

Then $(A4)_{3.5.1}$ with Proposition 3.3.4 implies that $|U_i(j)| \geq (1-2\varepsilon^{1/50})|U_i|$. For each $\vec{ij} \in E(D)$, we define

$$U_j^D(i) := \bigcap_{i' \in N_Q(i)} U_j(i'), \tag{3.5.4}$$

then we have

$$|U_j^D(i)| \geq (1 - 2(k-1)\varepsilon^{1/50})|U_j| \geq n/(2Tr). \tag{3.5.5}$$

In Step 4 we will map some vertices $x \in V(H)$ whose 'natural' image would have been in V_i to $U_j^D(i)$ instead, in order to 'balance out' the vertex class sizes.

Claim 3.5.1. *There exists a set $I^* = \{i_1^*, \ldots, i_k^*\} \subseteq [r]$ of k distinct numbers such that for any $k' \in [k]$ and $j \in [r]$, there exists a directed path $P(i_{k'}^*, j)$ from $i_{k'}^*$ to j in D.*

Proof. First, we claim that all $i \neq j \in [r]$ satisfy that $N_D^-(i) \cap N_D^-(j) \neq \emptyset$. Indeed, as $|N_R(\{i,j\})| \geq 2\delta(R) - r \geq (1 - 2/k + 2\sigma)r$, we have that

$$|\{s \in [r'] : |N_{R,V(Q_s)}(\{i,j\})| \geq k-1\}| \geq \sigma r \geq 3.$$

Thus there exists $s \in [r']$ such that $i, j \notin V(Q_s)$ while $|N_{R,V(Q_s)}(\{i,j\})| \geq k-1$. We choose $j' \in V(Q_s)$ such that $Q_s \setminus \{j'\} \subseteq N_R(\{i,j\})$, then (3.5.3) implies that $i, j \in N_D^+(j')$.

Now, we consider a number $i \in [r]$ which maximizes $|A(i)|$, where

$$A(i) = \{j \in [r] : \text{ there exists a directed path from } i \text{ to } j \text{ in } D\}.$$

If there exists $j \in [r]$ such that $j \notin A(i)$, then by the above claim, there exists $j' \in [r]$ such that $i, j \in N_D^+(j')$. Then $A(i) \cup \{j\} \subseteq A(j')$, which is a contradiction to the maximality of $A(i)$. Thus, we have $A(i) = [r]$. Let $i_1^* := i$.

Since $d_R(i_1^*) \geq \delta(R) \geq (1 - 1/k + \sigma)r$ by $(A5)_{3.5.1}$, we have $|\{s \in [r'] : N_{R,V(Q_s)}(i_1^*) = k\}| \geq \sigma r$. Thus, there exists $s \in [r']$ such that $V(Q_s) \subseteq N_R(i_1^*)$, and this with (3.5.3) implies that $V(Q_s) \subseteq N_D^-(i_1^*)$. We let $i_2^*, \ldots, i_k^*$ be $k-1$ arbitrary numbers in $V(Q_s)$. Then for all $k' \in [k]$ and $j \in [r]$, there exists a directed path from $i_{k'}^*$ to i_1^* and a directed path from i_1^* to j in D. Thus there exists a directed path from $i_{k'}^*$ to j in D. This proves the claim. □

We will now determine the approximate class sizes $\tilde{n}_i$ that our partition of H will have. For this, we first partition U_0 into $U_1', \ldots, U_r'$ in such a way that the vertices in U_i' are 'well connected' to the blow-up of the k-clique in Q to which i belongs.

$$\text{For all } i \in [r], u \in U'_i \text{ and } j \in N_Q(i), \text{ we have } d_{G,V_j}(u) \geq d^3 n' \text{ and } |U'_i| \leq 2\varepsilon^{3/4} n/r. \quad (3.5.6)$$

Indeed, it is easy to greedily construct such a partition by using the fact that $|U_0| \leq \varepsilon n$ and (A9)$_{3.5.1}$.

For $i \in I^*$, we will slightly increase the partition class sizes (cf. (3.5.9) and (X5)$_{3.5.1}$) as this will allow us to subsequently move any excess vertices from classes corresponding to I^* to another arbitrary class via the paths provided by Claim 3.5.1. For each $i \in [r]$, we let

$$n_i := n' + |U_i| + |U'_i| = |V_i| + |U_i| + |U'_i|, \quad (3.5.7)$$

then we have

$$n_i = (1 - 1/T \pm 2\varepsilon)n/r + (1 \pm 2\varepsilon)n/(Tr) \pm 2\varepsilon^{3/4} n/r = (1 \pm \varepsilon^{2/3}/2)n/r \text{ and } \sum_{i \in [r]} n_i = n - n_0. \quad (3.5.8)$$

For each $i \in [r]$ we let

$$\tilde{n}_i := \begin{cases} n_i + (r' - 1)\eta^{1/5} n & \text{if } i \in I^*, \\ n_i - \eta^{1/5} n & \text{if } i \in [r] \setminus I^*. \end{cases} \quad (3.5.9)$$

This with (3.5.8) implies that for each $i \in [r]$,

$$\tilde{n}_i = \frac{(1 \pm \varepsilon^{2/3})n}{r} \text{ and } \sum_{i \in [r]} \tilde{n}_i = \sum_{i \in [r]} n_i = n - n_0. \quad (3.5.10)$$

Step 2. Preparation of the graphs in $\mathcal{H}$. First, we will partition $\mathcal{H}$ into T collections $\mathcal{H}_1, \dots, \mathcal{H}_T$. Later we will pack each $\mathcal{H}_t$ into $G \cup F \cup \bigcup_{v \in V_0} F'_{v,t}$. (Recall that the $F'_{v,t}$ form a decomposition of F'.) As $G \cup F \cup F'$ has vertex partition $V_0, \dots, V_r, U_1, \dots, U_r, U'_1, \dots, U'_r$, for each $H \in \mathcal{H}$, we also need a suitable partition of $V(H)$ which is compatible with the partition of the host graph $G \cup F \cup F'$. To achieve this, we will apply Lemma 3.4.1 to each graph $H \in \mathcal{H}_t$ with the hypergraphs $\mathcal{C}_t$ and $\mathcal{C}^*_t$ to find the desired partition of $V(H)$.

By (3.5.1) we can partition $\mathcal{H}$ into $\mathcal{H}_1, \dots, \mathcal{H}_T$ such that for each $t \in [T]$,

$$\begin{aligned} e(\mathcal{H}_t) &= e(\mathcal{H})/T \pm \Delta n \overset{(\mathrm{A2})_{3.5.1}}{\leq} (1-2\nu/3)\alpha(k-1)n^2/(2Tr), \text{ and} \\ |\mathcal{H}_t| &\overset{(\mathrm{A1})_{3.5.1}}{\leq} 4e(\mathcal{H}_t)/n \leq 2\alpha(k-1)n/(Tr). \end{aligned} \tag{3.5.11}$$

For each $t \in [T]$, we wish to apply the randomised algorithm given by Lemma 3.4.1 with the following objects and parameters independently for all $H \in \mathcal{H}_t$.

object/parameter	H	R	Q	$\mathcal{C}_t$	$\mathcal{C}_t^*$	n_0	$C_{v_\ell,t}$	$C^*_{v_\ell,t}$	$\lceil 3/d \rceil$	η	ε	k	Δ	r	$\tilde{n}_i$
playing the role of	H	R	Q	$\mathcal{F}$	$\mathcal{F}^*$	n'	C_ℓ	C^*_ℓ	h	η	ε	k	Δ	r	n_i

Indeed, $(\mathrm{A5})_{3.5.1}$, $(\mathrm{A8})_{3.5.1}$ imply that $(\mathrm{A1})_{3.4.1}$, $(\mathrm{A2})_{3.4.1}$ and $(\mathrm{A3})_{3.4.1}$ hold with the above objects and parameters, respectively. Moreover, (3.5.10) implies that $(\mathrm{A4})_{3.4.1}$ holds too. Thus we obtain a partition $X_1^H, \dots, X_r^H$, $Y_1^H, \dots, Y_r^H$, $Z_r^H, \dots, Z_r^H, A^H$ of $V(H)$ such that $A^H = \{a_1^H, \dots, a_{n_0}^H\}$ is a 3-independent set of H and the following hold, where $X^H := \bigcup_{i\in[r]} X_i^H$, $Y^H := \bigcup_{i\in[r]} Y_i^H$, and $Z^H := \bigcup_{i\in[r]} Z_i^H$.

$(\mathrm{X1})_{3.5.1}$ For each $\ell \in [n_0]$, we have $d_H(a_\ell^H) \leq \frac{(2+d)e(H)}{n}$,

$(\mathrm{X2})_{3.5.1}$ for each $\ell \in [n_0]$, we have $N_H(a_\ell^H) \subseteq \bigcup_{i \in C_{v_\ell,t}} Y_i^H \setminus N_H^1(Z^H)$,

$(\mathrm{X3})_{3.5.1}$ $H[X^H]$ admits the vertex partition $(Q, X_1^H, \dots, X_r^H)$, and $H \setminus E(H[X^H])$ admits the vertex partition $(R, X_1^H \cup Y_1^H \cup Z_1^H, \dots, X_r^H \cup Y_r^H \cup Z_r^H)$,

$(\mathrm{X4})_{3.5.1}$ for each $ij \in E(Q)$, we have $e_H(X_i^H, X_j^H) = \frac{2e(H)\pm\varepsilon^{1/5}n}{(k-1)r}$,

$(\mathrm{X5})_{3.5.1}$ for each $i \in [r]$, we have $|Y_i^H| \leq 2\varepsilon^{1/3}n/r$ and $|X_i^H| + |Y_i^H| + |Z_i^H| = \tilde{n}_i \pm \eta^{1/4}n$; in particular, this with (3.5.9) implies that for each $i \in [r]$, we have

$$\widehat{n}_i^H := |X_i^H| + |Y_i^H| + |Z_i^H| \in \begin{cases} [n_i, n_i + \eta^{1/6}n] & \text{if } i \in I^*, \\ [n_i - \eta^{1/6}n, n_i] & \text{otherwise,} \end{cases}$$

$(\mathrm{X6})_{3.5.1}$ $N_H^1(X^H) \setminus X^H \subseteq Z^H$, and $|Z^H| \leq 4\Delta^{3k^3}\eta^{0.9}n$, (X7)$_{3.5.1}$ for all $\ell \in [n_0]$ and $i \in C_{v_\ell,t}$, we have $\mathbb{E}[N_H(a_\ell^H) \cap Y_i^H] \leq \frac{(2+d)e(H)}{(k-1)n}$.

By applying this randomised algorithm independently for each $H \in \mathcal{H}_1 \cup \cdots \cup \mathcal{H}_T$, we obtain that for all $t \in [T]$, $\ell \in [n_0]$ and $i \in C_{v_\ell,t}$, we have $\mathbb{E}[\sum_{H\in\mathcal{H}_t} |N_H(a_\ell^H) \cap Y_i^H|] \leq \frac{(2+d)e(\mathcal{H}_t)}{(k-1)n}$. Note that for each $H \in \mathcal{H}_t$, we have $|N_H(a_\ell^H) \cap Y_i^H| \leq \Delta$. As our applications of the randomised algorithm are independent for all $H \in \mathcal{H}_t$, a Chernoff bound (Lemma 3.3.1) together with (A2)$_{3.5.1}$ implies that for all $t \in [T]$, $\ell \in [n_0]$ and $i \in C_{v_\ell,t}$, we have

$$\mathbb{P}\Big[\sum_{H\in\mathcal{H}_t} |N_H(a_\ell^H) \cap Y_i^H| \geq \frac{2(1+d)e(\mathcal{H}_t)}{(k-1)n}\Big] \leq 2\exp\Big(-\frac{d^2 e(\mathcal{H}_t)^2/((k-1)^2n^2)}{2\Delta^2|\mathcal{H}_t|}\Big) \overset{(3.5.11),(A2)_{3.5.1}}{\leq} e^{-n^{1/3}}.$$

By taking a union bound over all $t \in [T], \ell \in [n_0]$ and $i \in C_{v_\ell,t}$, we can show that the following property (X8)$_{3.5.1}$ holds with probability at least $1 - kTn_0e^{-n^{1/3}} > 0$.

(X8)$_{3.5.1}$ For all $t \in [T]$, $\ell \in [n_0]$ and $i \in C_{v_\ell,t}$, we have $\sum_{H\in\mathcal{H}_t} |N_H(a_\ell^H) \cap Y_i^H| \leq \frac{2(1+d)e(\mathcal{H}_t)}{(k-1)n}$.

Thus we conclude that for all $H \in \mathcal{H}$ there exist partitions $X_1^H, \dots, X_r^H$, $Y_1^H, \dots, Y_r^H$, $Z_r^H, \dots, Z_r^H$, A^H of $V(H)$ such that $A^H = \{a_1^H, \dots, a_{n_0}^H\}$ is a 3-independent set of H and such that (X1)$_{3.5.1}$–(X6)$_{3.5.1}$ and (X8)$_{3.5.1}$ hold.

Note that $\sum_{i\in[r]} \widehat{n}_i^H = |V(H)| - |A^H| = n - n_0$. This with (3.5.8) implies that for each $H \in \mathcal{H}$, we have

$$\sum_{i\in I^*} (\widehat{n}_i^H - n_i) = \sum_{i\in[r]\setminus I^*} (n_i - \widehat{n}_i^H). \tag{3.5.12}$$

The following claim determines the number of vertices that we will redistribute via D.

Claim 3.5.2. *For each $H \in \mathcal{H}$, there exists a function $f^H : E(D) \to [\eta^{1/7}n] \cup \{0\}$ such that for each $i \in [r]$, we have*

$$\sum_{j\in N_D^+(i)} f^H(\vec{ij}) - \sum_{j\in N_D^-(i)} f^H(\vec{ji}) = \widehat{n}_i^H - n_i.$$

Proof. By (X5)$_{3.5.1}$, for each $i \in I^*$, we have $\widehat{n}_i^H - n_i \geq 0$ and for each $i \in [r] \setminus I^*$, we have $n_i - \widehat{n}_i^H \geq 0$. Thus by (3.5.12), there exists a bijection g^H from

$$\bigcup_{i\in I^*} \{i\} \times [\widehat{n}_i^H - n_i] \text{ to } \bigcup_{i\in[r]\setminus I^*} \{i\} \times [n_i - \widehat{n}_i^H].$$

For all $i \in I^*$ and $m \in [\widehat{n}_i^H - n_i]$, let $g^H(i,m) =: (g_1^H(i,m), g_2^H(i,m))$ and let $P_{i,m}$ be a directed path from i to $g_1^H(i,m)$ in D, which exists by Claim 3.5.1. As g^H is a bijection, for each $i \in [r]$, we have

$$|(g_1^H)^{-1}(i)| = \begin{cases} 0 & \text{if } i \in I^*, \\ n_i - \widehat{n}_i^H & \text{otherwise.} \end{cases} \tag{3.5.13}$$

For each $\vec{ij} \in E(D)$, we let

$$f^H(\vec{ij}) := |\{(i',m) : i' \in I^*, m \in [\widehat{n}_{i'}^H - n_{i'}] \text{ and } \vec{ij} \in E(P_{i',m})\}|.$$

Then for each $\vec{ij} \in E(D)$, we have

$$f^H(\vec{ij}) \leq \Big| \bigcup_{i' \in I^*} \{i'\} \times [\widehat{n}_{i'}^H - n_{i'}] \Big| \overset{(\text{X5})_{3.5.1}}{\leq} k\eta^{1/6} n \leq \eta^{1/7} n.$$

Note that for any $i \in I^*$ and $m \in [\widehat{n}_i^H - n_i]$, the path $P_{i,m}$ starts from a vertex in I^* and ends at $[r] \setminus I^*$. Thus for each $i \in [r]$ we have

$$\begin{aligned} &\sum_{j \in N_D^+(i)} f^H(\vec{ij}) - \sum_{j \in N_D^-(i)} f^H(\vec{ji}) \\ &= |\{(i',m) : m \in [\widehat{n}_{i'}^H - n_{i'}], i = i' \in I^*\}| - |\{(i',m) : i' \in I^*, m \in [\widehat{n}_{i'}^H - n_{i'}], g_1^H(i',m) = i\}| \\ &= \begin{cases} (\widehat{n}_i^H - n_i) - 0 = \widehat{n}_i^H - n_i & \text{if } i \in I^*, \\ 0 - (g_1^H)^{-1}(i) \overset{(3.5.13)}{=} \widehat{n}_i^H - n_i & \text{otherwise.} \end{cases} \end{aligned}$$

This proves the claim. □

For each $H \in \mathcal{H}$, we fix a function f^H satisfying Claim 3.5.2. For each $\vec{ij} \notin E(D)$, it will be convenient to set $f^H(\vec{ij}) := 0$.

We aim to embed vertices in $X_i^H \cup Y_i^H \cup Z_i^H$ into $V_i \cup U_i \cup U_i'$. As $|V_i \cup U_i \cup U_i'| = n_i$, by (3.5.7), it would be ideal if $|X_i^H \cup Y_i^H \cup Z_i^H| = n_i$ and $|X_i^H| = n'$. However, $(\text{X5})_{3.5.1}$ only guarantees that this is approximately true. In order to deal with this, we will use D and f^H to assign a small number of 'excess' vertices $u \in X_i^H$ into U_j when $\vec{ij} \in E(D)$. The definition of D will ensure that the image of u still has many neighbours in $V_{i'}$ for all $i' \in N_Q(i)$.

Step 3. Packing the graphs $H[Y^H \cup Z^H \cup A^H]$ into $G[U] \cup F'$. Now, we aim to find a suitable function ϕ' which packs $\{H[Y^H \cup Z^H \cup A^H] : H \in \mathcal{H}\}$ into $G[U] \cup F'$. In order to find ϕ', we will use Lemma 3.3.6. Moreover, we choose ϕ' in such a way that we can later extend ϕ' into a packing of the entire graphs $H \in \mathcal{H}$. One important property we need to ensure is the following: for any vertex $x \in X_j^H$ which is not embedded by ϕ', and any vertices $y_1, \ldots, y_i \in N_H(x) \cap (Y^H \cup Z^H)$ which are already embedded by ϕ', we need $N_G(\phi'(\{y_1, \ldots, y_i\})) \cap V_j$ to be large, so that x can be later embedded into $N_G(\phi'(\{y_1, \ldots, y_i\})) \cap V_j$. For this, we will introduce a hypergraph $\mathcal{N}_H$ which encodes information about the set $N_H(x) \cap (Y^H \cup Z^H)$ for each vertex $x \in X^H$. In order to describe the structure of G and H more succinctly, we also introduce a graph R' on $[2r]$ such that

$$E(R') = \{ij : (i-r)(j-r) \in E(R) \text{ or } i(j-r) \in E(R)\}.$$

For all $i \in [r]$ and $H \in \mathcal{H}$, let $V_{i+r} := U_i$ and $X^H_{i+r} := Y_i^H \cup Z_i^H$. Note that $(\mathrm{X3})_{3.5.1}$ and $(\mathrm{A4})_{3.5.1}$ imply that for each $H \in \mathcal{H}$,

$$\begin{aligned}&H[Y^H \cup Z^H] \textit{ admits the vertex partition } (R', \emptyset, \ldots, \emptyset, X^H_{r+1}, \ldots, X^H_{2r}), \textit{ and}\\ &G \textit{ is } (\varepsilon^{1/50}, (d^3))^+\textit{-regular with respect to the partition } (R', V_1, \ldots, V_{2r}).\end{aligned} \tag{3.5.14}$$

For all $H \in \mathcal{H}$ and $x \in X^H$, let

$$e_{H,x} := N_H(x) \setminus X^H \overset{(\mathrm{X6})_{3.5.1}}{=} N_H(x) \cap Z^H.$$

Let $\mathcal{N}_H$ be a multi-hypergraph on vertex set Z^H with

$$E(\mathcal{N}_H) := \{e_{H,x} : x \in N^1_H(Z^H) \cap X^H\}, \tag{3.5.15}$$

and let $f_H : E(\mathcal{N}_H) \to [r]$ be a function such that for all $x \in X^H$, we have that $x \in X^H_{f_H(e_{H,x})}$. Then $\Delta(\mathcal{N}_H) \leq \Delta$ and $\mathcal{N}_H$ has edge-multiplicity at most Δ. Note that, as $\mathcal{N}_H$ is a multi-hypergraph, there could be two distinct vertices $x \neq x' \in X^H$ such that $e_{H,x}$ and $e_{H,x'}$ consists of exactly the same vertices while $f_H(e_{H,x}) \neq f_H(e_{H,x'})$.

Our next aim is to construct a function ϕ' which packs $\{H[Y^H \cup Z^H \cup A^H] : H \in \mathcal{H}\}$ into $G[U] \cup F'$ in such a way that the following hold for all $H \in \mathcal{H}$.

$(\Phi'1)_{3.5.1}$ For each $e \in E(\mathcal{N}_H)$, we have $|N_G(\phi'(e)) \cap V_{f_H(e)}| \geq d^{5\Delta}|V_{f_H(e)}|$,

$(\Phi'2)_{3.5.1}$ for each $v \in V(G)$, we have $|\{H \in \mathcal{H} : v \in \phi'(H[Y^H \cup Z^H])\}| \leq \varepsilon^{1/8}n/r$,

$(\Phi'3)_{3.5.1}$ for all $i \in [r]$ and $H \in \mathcal{H}$, we have $\phi'(Y_i^H \cup Z_i^H) \subseteq U_i$, and

$(\Phi'4)_{3.5.1}$ $\phi'(A^H) = V_0$.

Claim 3.5.3. *There exists a function ϕ' packing $\{H[Y^H \cup Z^H \cup A^H] : H \in \mathcal{H}\}$ into $G[U] \cup F'$ which satisfies $(\Phi'1)_{3.5.1}$–$(\Phi'4)_{3.5.1}$.*

Proof. Let $\phi'_0 : \emptyset \to \emptyset$ be an empty packing. Let $H_1, \dots, H_\kappa$ be an enumeration of $\mathcal{H}$. For each $s \in [\kappa]$, let

$$\mathcal{H}^s := \{H_{s'}[Y^{H_{s'}} \cup Z^{H_{s'}} \cup A^{H_{s'}}] : s' \in [s]\}.$$

Our aim is to successively extend ϕ'_0 into $\phi'_1, \dots, \phi'_\kappa$ in such a way that each ϕ'_s satisfies the following.

$(\Phi'1)^s_{3.5.1}$ ϕ'_s packs $\mathcal{H}^s$ into $G[U] \cup F'$,

$(\Phi'2)^s_{3.5.1}$ for all $s' \in [s]$ and $e \in E(\mathcal{N}_{H_{s'}})$, we have $|N_G(\phi'_s(e)) \cap V_{f_{H_{s'}}(e)}| \geq d^{5\Delta}|V_{f_{H_{s'}}(e)}|$,

$(\Phi'3)^s_{3.5.1}$ for each $v \in V(G)$, we have $|\{s' \in [s] : v \in \phi'_s(H_{s'}[Y^{H_{s'}} \cup Z^{H_{s'}}])\}| \leq \varepsilon^{1/8}n/r$,

$(\Phi'4)^s_{3.5.1}$ for all $i \in [2r] \setminus [r]$ and $s' \in [s]$, we have $\phi'_s(X_i^{H_{s'}}) \subseteq V_i$,

$(\Phi'5)^s_{3.5.1}$ for all $s' \in [s]$ and $\ell \in [n_0]$, we have $\phi'_s(a_\ell^{H_{s'}}) = v_\ell$,

$(\Phi'6)^s_{3.5.1}$ for all $s' \in [s]$, $t \in [T]$ with $H_{s'} \in \mathcal{H}_t$, we have $\phi'_s(H_{s'}[Y^{H_{s'}} \cup Z^{H_{s'}} \cup A^{H_{s'}}]) \subseteq G[U] \cup \bigcup_{v \in V_0} F'_{v,t}$.

Note that ϕ'_0 vacuously satisfies $(\Phi'1)^0_{3.5.1}$–$(\Phi'6)^0_{3.5.1}$. Assume we have already constructed ϕ'_s satisfying $(\Phi'1)^s_{3.5.1}$–$(\Phi'6)^s_{3.5.1}$ for some $s \in [\kappa-1] \cup \{0\}$. We will show that we can construct ϕ'_{s+1}. Let

$$G(s) := G \setminus \phi'_s(\mathcal{H}^s).$$

For all $\ell \in [n_0]$ and $a_\ell^{H_{s+1}} \in A^{H_{s+1}}$, we first let

$$\psi(a_\ell^{H_{s+1}}) := v_\ell. \tag{3.5.16}$$

For each $i \in [2r] \setminus [r]$, let

$$V_i^{\mathrm{bad}} := \Big\{ v \in V_i : |\{s' \in [s] : v \in \phi'_{s'}(H_{s'}[Y^{H_{s'}} \cup Z^{H_{s'}}])\}| \geq \frac{\varepsilon^{1/8} n}{r} - 1 \Big\}.$$

Note that

$$|V_i^{\mathrm{bad}}| \overset{(\Phi'4)^s_{3.5.1}}{\leq} \frac{\sum_{s' \in [s]} |Y_{i-r}^{H_{s'}} \cup Z_{i-r}^{H_{s'}}|}{\frac{\varepsilon^{1/8} n}{r} - 1} \overset{(X5)_{3.5.1},(X6)_{3.5.1}}{\leq} 3\varepsilon^{1/3-1/8}\kappa \overset{(3.5.2)}{\leq} \frac{\varepsilon^{1/5} n}{r}. \tag{3.5.17}$$

Let $t \in [T]$ be such that $H_{s+1} \in \mathcal{H}_t$. For all $i \in [2r] \setminus [r]$ and $x \in X_i^{H_{s+1}}$, we let

$$B_x := \begin{cases} N_{F'_{v_\ell,t},V_i}(v_\ell) \setminus (N_{\phi'_s(\mathcal{H}^s)}(v_\ell) \cup V_i^{\mathrm{bad}}) & \text{if } x \in N_{H_{s+1}}(a_\ell^{H_{s+1}}) \cap X_i^{H_{s+1}} \text{ for some } \ell \in [n_0], \\ V_i \setminus V_i^{\mathrm{bad}} & \text{otherwise.} \end{cases}$$

We will later embed x into B_x. Note that if $x \in N_{H_{s+1}}(a_\ell^{H_{s+1}})$, then $x \notin N_{H_{s+1}}(a_{\ell'}^{H_{s+1}})$ for any $\ell' \in [n_0] \setminus \{\ell\}$ as $A^{H_{s+1}}$ is a 3-independent set in H_{s+1}. Also, if $x \in N_{H_{s+1}}(a_\ell^{H_{s+1}}) \cap X_i^{H_{s+1}}$, then by $(X2)_{3.5.1}$ we have $i - r \in C_{v_\ell,t}$. Thus in this case

$$\begin{aligned}
|B_x| &\geq d_{F'_{v_\ell,t},V_i}(v_\ell) - d_{\phi'_s(\mathcal{H}^s) \cap F'_{v_\ell,t},V_i}(v_\ell) - |V_i^{\mathrm{bad}}| \\
&\overset{(A7)_{3.5.1},(3.5.17)}{\geq} (1-d)\alpha|U_{i-r}| - d_{\phi'_s(\mathcal{H}^s) \cap F'_{v_\ell,t},V_i}(v_\ell) - \varepsilon^{1/5} n/r \\
&\overset{(X2)_{3.5.1},(\Phi'4)^s_{3.5.1},(\Phi'5)^s_{3.5.1},(\Phi'6)^s_{3.5.1}}{\geq} (1-d)\alpha|U_{i-r}| - \sum_{s' \in [s], H_{s'} \in \mathcal{H}_t} |N_{H_{s'}}(a_\ell^{H_{s'}}) \cap Y_{i-r}^{H_{s'}}| - \varepsilon^{1/5} n/r \\
&\overset{(X8)_{3.5.1}}{\geq} (1-d)\alpha|U_{i-r}| - \frac{2(1+d)e(\mathcal{H}_t)}{(k-1)n} - \varepsilon^{1/5} n/r \\
&\overset{(3.5.11)}{\geq} (1-d)\alpha|U_{i-r}| - \frac{(1+d)(1-2\nu/3)\alpha n}{Tr} - \varepsilon^{1/5} n/r \geq \alpha^2 |U_{i-r}| = \alpha^2 |V_i|.
\end{aligned}$$

If $x \notin N_{H_{s+1}}(a_\ell^{H_{s+1}})$ for any $\ell \in [n_0]$, then $|B_x| \geq |V_i| - |V_i^{\rm bad}| \geq (1-\varepsilon^{1/10})|V_i|$. So, for all $i \in [2r] \setminus [r]$ and $x \in X_i^{H_{s+1}}$, we have

$$B_x \subseteq V_i, \text{ and } |B_x| \geq \alpha^2 |V_i|. \tag{3.5.18}$$

For each $i \in [r]$, let $P_i := \emptyset$, and for each $i \in [2r] \setminus [r]$, let $P_i := X_i^{H_{s+1}}$. We wish to apply Lemma 3.3.6 with $H[Y^{H_{s+1}} \cup Z^{H_{s+1}}]$ playing the role of H and with the following objects and parameters.

object/parameter	$G(s)$	R'	V_i	P_i	$\varepsilon^{1/60}$	Δ	n'	α^2	d^3	$\mathcal{N}_{H_{s+1}}$	$f_{H_{s+1}}$	$1/(2T)$	B_x
playing the role of	G	R	V_i	X_i	ε	Δ	n	α	d	$\mathcal{M}$	f	β	A_x

Let us first check that we can indeed apply Lemma 3.3.6. Note that for each $ij \in E(R')$ with $i \in [2r] \setminus [r]$,

$$\begin{aligned} e_{G(s)}(V_i, V_j) &\geq e_G(V_i, V_j) - \Delta \sum_{v \in V_i} |\{s' \in [s] : v \in \phi'_s(H_{s'}[Y^{H_{s'}} \cup Z^{H_{s'}}])\}| \\ &\overset{(\Phi'3)^s_{3.5.1}}{\geq} e_G(V_i, V_j) - \Delta \varepsilon^{1/8} n |V_i|/r \overset{(A4)_{3.5.1}}{\geq} (1-\varepsilon^{1/9}) e_G(V_i, V_j). \end{aligned}$$

Thus (3.5.14) with Proposition 3.3.3 implies that $(A1)_{3.3.6}$ of Lemma 3.3.6 holds. Again (3.5.14) implies that $(A2)_{3.3.6}$ holds. Conditions $(A3)_{3.3.6}$ and $(A4)_{3.3.6}$ are obvious from $(A1)_{3.5.1}$, $(X3)_{3.5.1}$ and the definition of $\mathcal{N}_{H_{s+1}}$. Moreover, (3.5.18) implies that $(A5)_{3.3.6}$ also holds. Thus by Lemma 3.3.6, we obtain an embedding $\psi' : H_{s+1}[Y^{H_{s+1}} \cup Z^{H_{s+1}}] \to G(s)[U]$ satisfying the following.

$(P1)^{s+1}_{3.5.1}$ For each $x \in Y^{H_{s+1}} \cup Z^{H_{s+1}}$, we have $\psi'(x) \in B_x$,

$(P2)^{s+1}_{3.5.1}$ for each $e \in E(\mathcal{N}_{H_{s+1}})$, we have $|N_G(\psi'(e)) \cap V_{f_{H_{s+1}}(e)}| \geq (d^3/2)^{\Delta} |V_{f_{H_{s+1}}(e)}|$.

Let $\phi'_{s+1} := \phi_s \cup \psi \cup \psi'$. By (3.5.16) with the definitions of $G(s)$ and B_x, this implies $(\Phi'1)^{s+1}_{3.5.1}$ and $(\Phi'6)^{s+1}_{3.5.1}$. As $d \ll 1$, $(P2)^{s+1}_{3.5.1}$ implies $(\Phi'2)^{s+1}_{3.5.1}$, and the definitions of B_x and $V_i^{\rm bad}$ with $(P1)^{s+1}_{3.5.1}$ and $(\Phi'3)^s_{3.5.1}$ imply $(\Phi'3)^{s+1}_{3.5.1}$. Property $(P1)^{s+1}_{3.5.1}$ and (3.5.18) imply that $(\Phi'4)^{s+1}_{3.5.1}$

holds. $(\Phi'5)^{s+1}_{3.5.1}$ is obvious from (3.5.16). By repeating this for each $s \in [\kappa - 1]$, we can obtain our desired packing $\phi' := \phi'_\kappa$. Since $(\Phi'1)^{\kappa}_{3.5.1}$–$(\Phi'5)^{\kappa}_{3.5.1}$ imply that ϕ' is a packing of $\mathcal{H}^\kappa$ into $G[U] \cup F'$ satisfying $(\Phi'1)_{3.5.1}$–$(\Phi'4)_{3.5.1}$, this proves the claim. □

Step 4. Packing a 3-independent set $W^H \subseteq X^H$ into $U \cup U_0$. In the previous step, we constructed a function ϕ' packing $\{H[Y^H \cup Z^H \cup A^H] : H \in \mathcal{H}\}$ into $G[U] \cup F'$. However, for each graph $H \in \mathcal{H}$, the set $\phi'(H)$ only covers a small part of U. Eventually we need to cover every vertex of G with a vertex of H. Hence, for each $H \in \mathcal{H}$ we will choose a subset $W^H \subseteq X^H$ of size exactly $|U \cup U_0| - |Y^H \cup Z^H|$, and we will construct a function ϕ'' which packs $\{H[W^H] : H \in \mathcal{H}\}$ into $G[U \cup U_0]$. As later we will extend $\phi' \cup \phi''$ into a packing of $\mathcal{H}$ into $G \cup F \cup F'$, we again have to make sure that for any $x \in X_i^H \setminus W^H$ with neighbours in W^H, there is a sufficiently large set of candidates to which x can be embedded. In other words, the set $V_i \cap N(\phi''(N_H(x) \cap W^H))$ needs to be reasonably large. To achieve this, we choose W^H to be a 3-independent set, so $|N_H(x) \cap W^H| \leq 1$, and we will map each vertex $y \in N_H(x) \cap W^H$ into a vertex v which has a large neighbourhood in V_i.

Accordingly, for all $H \in \mathcal{H}$ and $i \in [r]$, we choose a subset $W_i^H \subseteq X_i^H$ satisfying the following:

$(\mathrm{W1})_{3.5.1}$ $\bigcup_{i \in [r]} W_i^H$ is a 3-independent set of H,

$(\mathrm{W2})_{3.5.1}$ for each $i \in [r]$, we have

$$|W_i^H| = |X_i^H| - n' \overset{(\mathrm{X5})_{3.5.1}}{=} n_i - n' - |Y_i^H| - |Z_i^H| \pm \eta^{1/6} n \overset{(3.5.7),(3.5.6),(\mathrm{X5})_{3.5.1}}{=} \frac{(1 \pm \varepsilon^{1/4})n}{Tr},$$

$(\mathrm{W3})_{3.5.1}$ $\bigcup_{i \in [r]} W_i^H \cap N_H^2(Z^H) = \emptyset$.

Indeed, the following claim ensures that there exist such sets W_i^H.

Claim 3.5.4. *For all $H \in \mathcal{H}$ and $i \in [r]$, there exists $W_i^H \subseteq X_i^H$ such that $(\mathrm{W1})_{3.5.1}$–$(\mathrm{W3})_{3.5.1}$ hold.*

Proof. We fix $H \in \mathcal{H}$. Assume that for some $i \in [r]$, we have already defined $W_1^H, \dots, W_{i-1}^H$ satisfying the following.

$(\mathrm{W}'1)^{i-1}_{3.5.1}$ $\bigcup_{i' \in [i-1]} W_{i'}^H$ is a 3-independent set of H,

$(\mathrm{W}'2)^{i-1}_{3.5.1}$ for each $i' \in [i-1]$, we have $|W_{i'}^H| = |X_{i'}^H| - n' = \frac{(1\pm\varepsilon^{1/4})n}{Tr}$,

$(\mathrm{W}'3)^{i-1}_{3.5.1}$ $\bigcup_{i' \in [i-1]} W_{i'}^H \cap N_H^2(Z^H) = \emptyset$.

Consider $W_i'^H := X_i^H \setminus (\bigcup_{i' \in [i-1]} N_H^2(W_{i'}^H) \cup N_H^2(Z^H))$. Note that $(\mathrm{X}6)_{3.5.1}$ implies that $|N_H^2(Z^H)| \leq 8\Delta^{3k^3+2}\eta^{0.9}n$. Also, $(\mathrm{X}3)_{3.5.1}$ with $(\mathrm{X}6)_{3.5.1}$ implies that

$$\bigcup_{i' \in [i-1]} N_H^2(W_{i'}^H) \cap X_i^H \subseteq N_H^1(Z^H) \cup \bigcup_{i' \in N_Q(i) \cap [i-1]} N_H^2(W_{i'}^H).$$

Thus

$$\begin{aligned} |W_i'^H| &\geq |X_i^H| - |N_H^2(Z^H)| - \sum_{i' \in N_Q(i)\cap[i-1]} |N_H^2(W_{i'}^H)| \\ &\overset{(\mathrm{W}'2)^{i-1}_{3.5.1}}{\geq} |X_i^H| - 8\Delta^{3k^3+2}\eta^{0.9}n - \frac{2k\Delta^2 n}{Tr} \overset{(\mathrm{X}5)_{3.5.1},(3.5.10)}{\geq} \Delta^3(|X_i^H| - n'). \end{aligned}$$

Thus, by Lemma 3.3.21, $W_i'^H$ contains a 3-independent set W_i^H of size $|X_i^H| - n'$. Then, by the choice of W_i^H, $(\mathrm{W}'1)^{i}_{3.5.1}$–$(\mathrm{W}'3)^{i}_{3.5.1}$ hold. By repeating this for all $i \in [r]$ in increasing order, we obtain W_i^H satisfying $(\mathrm{W}'1)^{r}_{3.5.1}$–$(\mathrm{W}'3)^{r}_{3.5.1}$, and thus satisfying $(\mathrm{W}1)_{3.5.1}$–$(\mathrm{W}3)_{3.5.1}$. This proves the claim. □

For all $H \in \mathcal{H}$ and $i \in [r]$, let $W^H := \bigcup_{i' \in [r]} W_{i'}^H$ and $W_i := \bigcup_{H \in \mathcal{H}} W_i^H$, where we consider the sets $V(H)$ to be disjoint for different $H \in \mathcal{H}$. Note that for all $H \in \mathcal{H}$ and $i \in [r]$, Claim 3.5.2 implies that $0 \leq \sum_{j \in N_D^+(i)} f^H(\vec{ij}) \leq r\eta^{1/7}n$. For all $H \in \mathcal{H}$ and $i \in [r]$, we choose a partition $W_i^{H,F}, W_i^{H,U'}, W_i^{H,D}$ of W_i^H such that

$$|W_i^{H,U'}| = |U_i'| \text{ and } |W_i^{H,D}| = \sum_{j \in N_D^+(i)} f^H(\vec{ij}) \leq r\eta^{1/7}n. \tag{3.5.19}$$

Such partitions exist by (3.5.6), $(\mathrm{W}2)_{3.5.1}$ and the fact that $\eta \ll \varepsilon \ll 1/T$. For each $S \in \{F, D, U'\}$, we let $W^{H,S} := \bigcup_{i \in [r]} W_i^{H,S}$.

We now construct a function ϕ'' which maps all the vertices of W^H into $U_0 \cup (U \setminus \phi'(Y^H \cup Z^H))$ for each $H \in \mathcal{H}$. (In Step 6 we will then apply Theorem 3.3.15 to embed all the vertices of $X^H \setminus W^H$ into V.) We will define ϕ'' separately for $W^{H,F}, W^{H,D}$ and $W^{H,U'}$. We first cover the 'exceptional' set U_0 with $W^{H,U'}$. (3.5.19) implies that for all $H \in \mathcal{H}$ and $i \in [r]$, there exists a bijection $\phi''^{H}_{U',i}$ from $W_i^{H,U'}$ to U'_i. We let $\phi''_{U'} := \bigcup_{H\in\mathcal{H}} \bigcup_{i\in[r]} \phi''^{H}_{U',i}$. Then (3.5.6) implies the following.

$$\text{For all } i \in [r] \text{ and } H \in \mathcal{H}, \text{ the function } \phi''_{U'} \text{ is bijective between } W_i^{H,U'} \text{ and } U'_i. \text{ Moreover, for all } x \in W_i^{H,U'} \text{ and } j \in N_Q(i), \text{ we have } d_{G,V_j}(\phi''_{U'}(x)) \geq d^3 n'. \tag{3.5.20}$$

We intend to embed the neighbours of W_i^H into $\bigcup_{j\in N_Q(i)} V_j$. Thus it is natural to embed W_i^H into U_i and make use of (A6)$_{3.5.1}$. This is in fact what we will do for $W_i^{H,F}$. However, the vertices of $W_i^{H,D}$ will first be mapped to a suitable set of vertices in $U_j^D(i) \subseteq U_j$ for $j \in N_D^+(i)$. The definition of D and f^H will ensure that the remaining uncovered part of each U_j matches up exactly with the size of each $W_j^{H,F}$.

By (3.5.5), for all $\vec{ij} \in E(D)$ and $H \in \mathcal{H}$, we have

$$|U_j^D(i) \setminus \phi'(Y^H \cup Z^H)| \geq n/(2Tr) - |Y_j^H \cup Z_j^H| \overset{(X5)_{3.5.1},(X6)_{3.5.1}}{\geq} |U_j|/3.$$

For $i \in [r]$ and $H \in \mathcal{H}$, we let

$$b_i^H := \sum_{j\in N_D^-(i)} f^H(\vec{ji}) \overset{\text{Claim 3.5.2}}{\leq} r\eta^{1/7} n \leq \eta^{1/10}|U_i|.$$

Thus, for each $i \in [r]$, we can apply Lemma 3.3.18 with the following objects and parameters.

object/parameter	κ	r	$H \in \mathcal{H}$	U_i	$j \in [r]$	$U_i^D(j) \setminus \phi'(Y^H \cup Z^H)$	$\eta^{1/10}$	$f^H(\vec{ji})$	b_i^H	$1/3$
playing the role of	s	r	$i \in [s]$	A	$j \in [r]$	$A_{i,j}$	ε	$m_{i,j}$	$\sum_{j\in[r]} m_{i,j}$	d

(Recall that $f^H(\vec{ji}) = 0$ if $\vec{ji} \notin E(D)$.) Then we obtain sets $U_{i,j}^H \subseteq U_i$ satisfying the following for each $i \in [r]$, where $U_i^H := \bigcup_{j\in[r]} U_{i,j}^H$.

(U1)$_{3.5.1}$ For each $j \in [r]$ and $H \in \mathcal{H}$, we have $|U_{i,j}^H| = f^H(\vec{ji})$ and $U_{i,j}^H \subseteq U_i^D(j) \setminus \phi'(Y^H \cup Z^H)$, (U2)$_{3.5.1}$ for $j \neq j' \in [r]$ and $H \in \mathcal{H}$, we have $U^H_{i,j} \cap U^H_{i,j'} = \emptyset$,

(U3)$_{3.5.1}$ for each $v \in U_i$, we have $|\{H \in \mathcal{H} : v \in U^H_i\}| \leq \eta^{1/20}|\mathcal{H}| \overset{(3.5.2)}{\leq} \eta^{1/20} n$.

Now for all $H \in \mathcal{H}$ and $i \in [r]$, we partition $W^{H,D}_i$ into $W^{H,D}_{i,1}, \dots, W^{H,D}_{i,r}$ in such a way that $|W^{H,D}_{i,j}| = f^H(\vec{ij})$. Clearly, this is possible by (3.5.19). Thus (U1)$_{3.5.1}$ implies that for all $(i,j) \in [r] \times [r]$ and $H \in \mathcal{H}$, we have $|W^{H,D}_{i,j}| = f^H(\vec{ij}) = |U^H_{j,i}|$. Thus there exists a bijection $\phi''^H_{D,i,j} : W^{H,D}_{i,j} \to U^H_{j,i}$. Let $\phi''_D := \bigcup_{(i,j)\in[r]\times[r]} \bigcup_{H\in\mathcal{H}} \phi''^H_{D,i,j}$. Then, for $\vec{ij} \in E(D), H \in \mathcal{H}$ and $y \in W^{H,D}_{i,j}$, we have that

$$\phi''_D(y) \in U^H_{j,i} \subseteq U^D_j(i) \setminus \phi'(Y^H \cup Z^H).$$

Thus, (3.5.4) with (U1)$_{3.5.1}$ and (U2)$_{3.5.1}$ implies the following.

For each $H \in \mathcal{H}$, the function ϕ''_D is bijective between $\bigcup_{i\in[r]} W^{H,D}_i = W^{H,D}$ and $\bigcup_{i\in[r]} U^H_i$. Moreover, for all $x \in W^{H,D}_i$ and $j' \in N_Q(i)$, we have $d_{G,V_{j'}}(\phi''_D(x)) \geq d^3 n'/2$. (3.5.21)

Now, for all $H \in \mathcal{H}$ and $i \in [r]$

$$\begin{aligned}
|W^{H,F}_i| &= |W^H_i| - |W^{H,U'}_i| - |W^{H,D}_i| \overset{(3.5.19),(W2)_{3.5.1}}{=} (|X^H_i| - n') - |U'_i| - \sum_{j\in N^+_D(i)} f^H(\vec{ij}) \\
&\overset{(X5)_{3.5.1}}{=} \widehat{n}^H_i - \sum_{j\in N^+_D(i)} f^H(\vec{ij}) - |Y^H_i| - |Z^H_i| - n' - |U'_i| \\
&\overset{\text{Claim 3.5.2}}{=} n_i - \sum_{j\in N^-_D(i)} f^H(\vec{ji}) - |Y^H_i| - |Z^H_i| - n' - |U'_i| \\
&\overset{(3.5.7),(U1)_{3.5.1}}{=} |U_i| - |Y^H_i| - |Z^H_i| - \sum_{j\in N^-_D(i)} |U^H_{i,j}| \overset{(\Phi'3)_{3.5.1}}{=} |U_i \setminus (\phi'(Y^H_i \cup Z^H_i) \cup U^H_i)|.
\end{aligned}$$

Thus, there exists a bijection $\phi''^H_{F,i}$ from $W^{H,F}_i$ to $U_i \setminus (\phi'(Y^H_i \cup Z^H_i) \cup U^H_i)$. Let $\phi''_F := \bigcup_{H\in\mathcal{H}} \bigcup_{i\in[r]} \phi''^H_{F,i}$. Then (A6)$_{3.5.1}$ implies the following.

For all $H \in \mathcal{H}$ and $i \in [r]$, the function ϕ''_F is bijective between $W^{H,F}_i$ and $U_i \setminus (\phi'(Y^H_i \cup Z^H_i) \cup U^H_i)$. Moreover, for all $x \in W^{H,F}_i$ and $j \in N_Q(i)$, we have $d_{F,V_j}(\phi''_F(x)) \geq d^3 n'$. (3.5.22)

We define

$$\phi'' := \phi''_{U'} \cup \phi''_D \cup \phi''_F \text{ and } \phi_* := \phi' \cup \phi''. \tag{3.5.23}$$

Then (3.5.20), (3.5.21) and (3.5.22) imply that ϕ'' is bijective between W^H and $(U \cup U_0) \setminus \phi'(Y^H \cup Z^H)$, when restricted to W^H for each $H \in \mathcal{H}$. Thus, we know that

$$\phi_* \text{ is bijective between } W^H \cup Y^H \cup Z^H \cup A^H \text{ and } U \cup U_0 \cup V_0 \text{ for each } H \in \mathcal{H}. \tag{3.5.24}$$

Moreover, (3.5.20), (3.5.21) and (3.5.22) imply that the following hold for all $i \in [r]$ and $H \in \mathcal{H}$.

$(\Phi_*1)_{3.5.1}$ If $x \in W_i^{H,F}$, then $\phi_*(x) \in U$ and, for each $j \in N_Q(i)$, we have $d_{F,V_j}(\phi_*(x)) \geq d^3 n'$,

$(\Phi_*2)_{3.5.1}$ if $x \in W_i^{H,D}$, then $\phi_*(x) \in U$ and, for each $j \in N_Q(i)$, we have $d_{G,V_j}(\phi_*(x)) \geq d^3 n'/2$,

$(\Phi_*3)_{3.5.1}$ if $x \in W_i^{H,U'}$, then $\phi_*(x) \in U_0$ and, for each $j \in N_Q(i)$, we have $d_{G,V_j}(\phi_*(x)) \geq d^3 n'$.

Furthermore, $(\Phi'2)_{3.5.1}$ with $(U3)_{3.5.1}$ implies that

$(\Phi_*4)_{3.5.1}$ for $u \in U$, we have $|\{H \in \mathcal{H} : u \in \phi_*(Y^H \cup Z^H \cup W^{H,D})\}| \leq 2\varepsilon^{1/8} n/r$.

Step 5. Packing the graphs $H[X^H \setminus W^H]$ into internally regular graphs. Note that $(X6)_{3.5.1}$ and $(W3)_{3.5.1}$ together imply that $N_H(W^H) \cap (Y^H \cup Z^H \cup A^H) = \emptyset$ for each $H \in \mathcal{H}$. This implies that ϕ_* is a function packing $\{H[Y^H \cup Z^H \cup W^H \cup A^H] : H \in \mathcal{H}\}$ into $G[U \cup U_0] \cup F'$. We wish to pack the remaining part $H[X^H \setminus W^H]$ of each $H \in \mathcal{H}$ into $G[V]$ by using Theorem 3.3.15. In order to be able to apply Theorem 3.3.15, we first need to pack suitable subcollections of $\mathcal{H}$ into internally q-regular graphs. More precisely, for each $t \in [T]$, we will partition $\mathcal{H}_t$ into $\mathcal{H}_{t,1}, \dots, \mathcal{H}_{t,w}$ and apply Lemma 3.3.14 to the unembedded part of each graph in $\mathcal{H}_{t,w'}$ to pack all these parts into a graph $H_{t,w'}$ on $|V|$ vertices which is internally q-regular. We can then use Theorem 3.3.15 to pack all the $H_{t,w'}$ into $G[V]$ in Step 6.

For this purpose, we choose an integer q and a constant ξ such that $1/T \ll 1/q \ll \xi \ll \alpha$ and let

$$w := \frac{e(\mathcal{H})}{(1-3\xi)T(k-1)qn/2} \overset{(A2)_{3.5.1}}{\leq} \frac{(1-\nu/2)\alpha n'}{qT}. \tag{3.5.25}$$

By using (3.5.1) and (3.5.11), for each $t \in [T]$, we can further partition $\mathcal{H}_t$ into $\mathcal{H}_{t,1}, \dots, \mathcal{H}_{t,w}$ such that for each $(t, w') \in [T] \times [w]$, we have

$$e(\mathcal{H}_{t,w'}) = (1-3\xi)(k-1)qn/2 \pm 2\Delta n = (1-3\xi \pm \xi/2)(k-1)qn/2. \tag{3.5.26}$$

By $(A1)_{3.5.1}$, we have

$$|\mathcal{H}_{t,w'}| \leq 2(k-1)q \leq (q\xi)^{3/2}. \tag{3.5.27}$$

For all $H \in \mathcal{H}$ and $i \in [r]$, let $\widetilde{X}_i^H := X_i^H \setminus W_i^H$ and $\widetilde{X}^H := \bigcup_{i\in[r]} \widetilde{X}_i^H$. Thus, by $(W2)_{3.5.1}$ we have $|\widetilde{X}_i^H| = n'$ for all $H \in \mathcal{H}$ and $i \in [r]$. Moreover, for all $t \in [T]$, $w' \in [w]$ and $ij \in E(Q)$, we have

$$\begin{aligned}\sum_{H\in\mathcal{H}_{t,w'}} e(H[\widetilde{X}_i^H, \widetilde{X}_j^H]) &= \sum_{H\in\mathcal{H}_{t,w'}} (e(H[X_i^H, X_j^H]) \pm \Delta(|W_i^H| + |W_j^H|) \\ &\overset{(X4)_{3.5.1},(W2)_{3.5.1}}{=} \sum_{H\in\mathcal{H}_{t,w'}} \left(\frac{2e(H) \pm \varepsilon^{1/5}n}{(k-1)r} \pm \frac{3\Delta n}{Tr}\right) \overset{(3.5.26)}{=} (1-3\xi \pm \xi)qn'. \end{aligned} \tag{3.5.28}$$

When packing $H[\tilde{X}^H]$ and $H'[\tilde{X}^{H'}]$ (say) into the same graph $H_{t,w'}$, we need to make sure that the 'attachment sets' of $H[\tilde{X}^H]$ and $H'[\tilde{X}^{H'}]$ are not mapped to the same vertex sets in $H_{t,w'}$. The attachment set for $H[\tilde{X}^H]$ contains those vertices of $\tilde{X}^H$ which have a neighbour in $W^H \cup Y^H \cup Z^H \cup A^H$ (more precisely, a neighbour in $W^H \cup Z^H$) and is defined in (3.5.29). Keeping these attachment sets disjoint in $H_{t,w'}$ ensures that we can make the embedding of each $\tilde{X}^H$ consistent with the existing partial embedding of H without attempting to use an edge of F or G twice. For all $i \in [r]$ and $H \in \mathcal{H}$, we let

$$N_i^{H,F} := \bigcup_{i'\in N_Q(i)} N_H^1(W_{i'}^{H,F}) \cap \widetilde{X}_i^H \quad \text{and } N_i^{H,G} := N_H^1(Z^H \cup W^{H,D} \cup \bigcup_{i'\in N_Q(i)} W_{i'}^{H,U'}) \cap \widetilde{X}_i^H. \tag{3.5.29}$$

Note that (W1)$_{3.5.1}$, (W3)$_{3.5.1}$ and the fact that $W^{H,F}, W^{H,D}, W^{H,U'}$ form a partition of W^H implies that

$$N_i^{H,F} \cap N_i^{H,G} = \emptyset. \tag{3.5.30}$$

Moreover, if $x \in N_i^{H,F}$ then x has a unique neighbour in $W^{H,F}$. Similarly, if $x \in N_i^{H,G}$, then either x has a unique neighbour in $W^{H,D} \cup W^{H,U'}$ or x has at least one neighbour in Z^H (but not both). Note that for $i \in [r]$ and $H \in \mathcal{H}$,

$$\begin{aligned} |N_i^{H,F} \cup N_i^{H,G}| &\leq \sum_{i' \in N_Q(i)} \Delta(|W_{i'}^{H,F}| + |W_{i'}^{H,U'}|) + \Delta(|Z^H| + |W^{H,D}|) \\ &\overset{\substack{(X6)_{3.5.1},\\ (W2)_{3.5.1},(3.5.19)}}{\leq} \frac{2\Delta kn}{Tr} + 4\Delta^{3k^3+1}\eta^{0.9}n + \Delta r^2 \eta^{1/7} n \leq T^{-2/3}n'. \end{aligned} \tag{3.5.31}$$

For each $i \in [r]$, we consider a set $\widehat{X}_i$ with $|\widehat{X}_i| = n'$ such that $\widehat{X}_1, \dots, \widehat{X}_r$ are pairwise vertex-disjoint. For each $(t, w') \in [T] \times [w]$, let $\mathcal{H}_{t,w'} =: \{H^1_{t,w'}, \dots, H^{h(t,w')}_{t,w'}\}$. Then, by (3.5.27), (3.5.28), (3.5.31) and (X3)$_{3.5.1}$, we can apply Lemma 3.3.14 with the following objects and parameters for each $(t, w') \in [T] \times [w]$.

object/parameter	$H^j_{t,w'}[\widetilde{X}^{H^j_{t,w'}}]$	$\widetilde{X}_i^{H^j_{t,w'}}$	$\widehat{X}_i$	n'	q	ξ	$T^{-2/3}$	$h(t,w')$	$N_i^{H^j_{t,w'},F} \cup N_i^{H^j_{t,w'},G}$	Q
playing the role of	L_j	X_i^j	V_i	n	q	ξ	ε	s	W_i^j	R

Then for each $(t, w') \in [T] \times [w]$, we obtain a function $\Phi_{t,w'}$ packing $\{H[\widetilde{X}^H] : H \in \mathcal{H}_{t,w'}\}$ into some graph $H_{t,w'}$ which is internally q-regular with respect to the vertex partition $(Q, \widehat{X}_1, \dots, \widehat{X}_r)$. Moreover, for all $i \in [r]$ and $H \in \mathcal{H}_{t,w'}$ we have $\Phi_{t,w'}(\widetilde{X}_i^H) = \widehat{X}_i$ and for distinct $H, H' \in \mathcal{H}_{t,w'}$ and $i \in [r]$, we have

$$\Phi_{t,w'}(N_i^{H,F} \cup N_i^{H,G}) \cap \Phi_{t,w'}(N_i^{H',F} \cup N_i^{H',G}) = \emptyset. \tag{3.5.32}$$

Note that for all $(t, w') \in [T] \times [w]$, the graphs $H_{t,w'}$ have same vertex set $\bigcup_{i\in[r]} \widehat{X}_i$. For all

$i \in [r]$ and $(t, w') \in [T] \times [w]$, we let

$$L_i^{t,w'} := \bigcup_{H \in \mathcal{H}_{t,w'}} \Phi_{t,w'}(N_i^{H,F}) \quad \text{and} \quad M_i^{t,w'} := \bigcup_{H \in \mathcal{H}_{t,w'}} \Phi_{t,w'}(N_i^{H,G}). \tag{3.5.33}$$

Then by (3.5.30) and (3.5.32) we have

$$L_i^{t,w'} \cup M_i^{t,w'} \subseteq \widehat{X}_i \quad \text{and} \quad L_i^{t,w'} \cap M_i^{t,w'} = \emptyset. \tag{3.5.34}$$

By (3.5.27) and (3.5.31), for all $(t, w') \in [T] \times [w]$ and $i \in [r]$

$$|L_i^{t,w'} \cup M_i^{t,w'}| \leq q^{3/2} T^{-2/3} n' \leq T^{-1/2} n'. \tag{3.5.35}$$

Step 6. Packing the internally regular graphs $H_{t,w'}$ into $G[V]$. In the previous step, we constructed a collection $\widehat{\mathcal{H}} := \{H_{1,1}, \ldots, H_{T,w}\}$ of internally q-regular graphs on $|V|$ vertices. We now wish to apply Theorem 3.3.15 to pack $\widehat{\mathcal{H}}$ into $G[V]$. However, our packing needs to be consistent with the packing ϕ_*. Note that for each $H \in \mathcal{H}$ the set $W^H \cup Y^H \cup Z^H \cup A^H$ consists of exactly those vertices of H which are already embedded by ϕ_*. Thus by (X3)$_{3.5.1}$, (X6)$_{3.5.1}$, (3.5.29) and (3.5.33), it follows that whenever $x \in \widehat{X}_i$ is a vertex of $H_{t,w'}$ such that the set $\Phi_{t,w'}^{-1}(x)$ of pre-images of x contains a neighbour of some vertex which is already embedded by ϕ_*, then $x \in L_i^{t,w'} \cup M_i^{t,w'}$. Thus in order to ensure that our packing of $\widehat{\mathcal{H}}$ is consistent with ϕ_*, for each $i \in [r]$, each $(t, w') \in [T] \times [w]$ and each $y \in L_i^{t,w'} \cup M_i^{t,w'}$ we will choose a suitable target set $A_y^{t,w'}$ of vertices of $G[V]$ and will map y into this set.

For all $(t, w') \in [T] \times [w]$, $i \in [r]$ and any vertex $y \in L_i^{t,w'} \cup M_i^{t,w'}$, (3.5.32) implies that there exists a unique graph $H_y^{t,w'} \in \mathcal{H}_{t,w'}$ and a unique vertex $x_y^{t,w'} \in N_i^{H_y^{t,w'},F} \cup N_i^{H_y^{t,w'},G}$ such that $y = \Phi_{t,w'}(x_y^{t,w'})$. Let

$$J_y^{t,w'} := N_{H_y^{t,w'}}(x_y^{t,w'}) \cap (W^{H_y^{t,w'}} \cup Z^{H_y^{t,w'}}) = N_{H_y^{t,w'}}(x_y^{t,w'}) \cap (W^{H_y^{t,w'}} \cup Y^{H_y^{t,w'}} \cup Z^{H_y^{t,w'}} \cup A^{H_y^{t,w'}}).$$

The final equality follows from (X6)$_{3.5.1}$. For all $(t, w') \in [T] \times [w]$, $i \in [r]$ and any vertex $y \in L_i^{t,w'} \cup M_i^{t,w'}$, we define the target set

$$A_y^{t,w'} := \begin{cases} N_F(\phi_*(J_y^{t,w'})) \cap V_i & \text{if } x_y^{t,w'} \in N_i^{H_y^{t,w'},F}, \\ N_G(\phi_*(J_y^{t,w'})) \cap V_i & \text{if } x_y^{t,w'} \in N_i^{H_y^{t,w'},G}. \end{cases}$$

Note that $A_y^{t,w'}$ is well-defined as (3.5.30) implies that exactly one of the above cases holds. Moreover, the following claim implies that these target sets are sufficiently large.

Claim 3.5.5. *For all $(t, w') \in [T] \times [w]$, $i \in [r]$ and any vertex $y \in L_i^{t,w'} \cup M_i^{t,w'}$, we have*

$$|A_y^{t,w'}| \geq d^{5\Delta}|V_i|.$$

Proof. We fix $(t, w') \in [T] \times [w]$, $i \in [r]$ and a vertex $y \in L_i^{t,w'} \cup M_i^{t,w'}$. For simplicity, we write $H := H_y^{t,w'}$, $x := x_y^{t,w'}$ and $J := J_y^{t,w'}$. Then (3.5.30) implies that exactly one of the following two cases holds.

Case 1. $x \in N_i^{H,F}$. In this case, $(\mathrm{W1})_{3.5.1}$ and $(\mathrm{W3})_{3.5.1}$ imply that

$$J = N_H(x) \cap W^{H,F} \overset{(\mathrm{X3})_{3.5.1}}{=} N_H(x) \cap \bigcup_{i' \in N_Q(i)} W_{i'}^{H,F} \quad \text{and} \quad |J| = 1.$$

Then by $(\Phi_*1)_{3.5.1}$, we know $|A_y^{t,w'}| \geq d^3|V_i|$.

Case 2. $x \in N_i^{H,G}$. In this case, by (3.5.29) and $(\mathrm{W3})_{3.5.1}$, we have exactly one of the following cases.

Case 2.1 $x \in N_H^1(Z^H)$. In this case, $N_H(x) \cap W^H = \emptyset$ by $(\mathrm{W3})_{3.5.1}$. Thus we have $J = N_H(x) \cap Z^H$. Then (3.5.15) and $(\Phi'1)_{3.5.1}$ imply that $|A_y^{t,w'}| = |N_G(\phi'(e_{H,x})) \cap V_{f_H(e_{H,x})}| \geq d^{5\Delta}|V_i|$.

Case 2.2 $x \in N_H^1(W^{H,D} \cup W^{H,U'})$. In this case, again $(\mathrm{W1})_{3.5.1}$, $(\mathrm{W3})_{3.5.1}$ and $(\mathrm{X3})_{3.5.1}$ imply that

$$J = N_H(x) \cap (W^{H,D} \cup W^{H,U'}) = N_H(x) \cap \bigcup_{i' \in N_Q(i)} (W_{i'}^{H,D} \cup W_{i'}^{H,U'}) \quad \text{and} \quad |J| = 1.$$

Thus $(\Phi_*2)_{3.5.1}$ or $(\Phi_*3)_{3.5.1}$ imply that $|A_y^{t,w'}| \geq d^3|V_i|/2$. This proves the claim. $\square$

Let $\mathbf{S} := [T] \times [w]$. Let Λ be the graph with

$$V(\Lambda) := \{(\vec{s}, y) : \vec{s} \in \mathbf{S}, y \in \bigcup_{\vec{s} \in \mathbf{S}, i \in [r]} L_i^{\vec{s}} \cup M_i^{\vec{s}}\}$$

and

$$E(\Lambda) := \left\{(\vec{s}, y)(\vec{t}, y') : \vec{s} \neq \vec{t} \in \mathbf{S}, i \in [r], (y, y') \in (L_i^{\vec{s}} \times L_i^{\vec{t}}) \cup (M_i^{\vec{s}} \times M_i^{\vec{t}}) \text{ and } \phi_*(J_y^{\vec{s}}) \cap \phi_*(J_{y'}^{\vec{t}}) \neq \emptyset\right\}.$$

Note that Λ is the graph indicating possible overlaps of images of distinct edges when we extend ϕ_*. Indeed, if $(\vec{s}, y)$ and $(\vec{t}, y')$ are adjacent in Λ, there are $z \in N_{H^{\vec{s}}_y}(x^{\vec{s}}_y)$ and $z' \in N_{H^{\vec{t}}_{y'}}(x^{\vec{t}}_{y'})$ such that $\phi_*(z) = \phi_*(z')$. If we embed y and y' onto the same vertex, then the two edges $x^{\vec{s}}_y z$ and $x^{\vec{t}}_{y'} z'$ would be embedded onto the same edge of $G \cup F$. Thus we need to ensure that $\phi(y) \neq \phi(y')$.

Note that for all $(\vec{s}, y) \in V(\Lambda)$ and $\vec{t} \in \mathbf{S}$, we have

$$\begin{aligned} |\{(\vec{t}, y') \in N_\Lambda((\vec{s}, y))\}| &\leq |\{y' : H^{\vec{t}}_{y'} \in \mathcal{H}_{\vec{t}}, \phi_*(J^{\vec{t}}_{y'}) \cap \phi_*(J^{\vec{s}}_y) \neq \emptyset\}| \\ &\leq \sum_{v \in \phi_*(J^{\vec{s}}_y)} |\{y' : H^{\vec{t}}_{y'} \in \mathcal{H}_{\vec{t}}, v \in \phi_*(J^{\vec{t}}_{y'})\}| \\ &\leq \sum_{v \in \phi_*(J^{\vec{s}}_y)} \sum_{H \in \mathcal{H}_{\vec{t}}} |\{x \in V(H) : v \in \phi_*(N_H(x))\}| \\ &\leq \sum_{v \in \phi_*(J^{\vec{s}}_y)} \sum_{H \in \mathcal{H}_{\vec{t}}} |\{x \in N_H(x') : v = \phi_*(x'), x' \in V(H)\}| \\ &\overset{(3.5.24)}{\leq} \sum_{v \in \phi_*(J^{\vec{s}}_y)} \sum_{H \in \mathcal{H}_{\vec{t}}} \Delta \leq \Delta^2 |\mathcal{H}_{\vec{t}}| \overset{(3.5.27)}{\leq} \Delta^2 (\xi q)^{3/2} \leq q^2. \end{aligned} \quad (3.5.36)$$

(Here the third inequality holds by the definition of $J^{\vec{t}}_{y'}$ and the definition of $x^{\vec{t}}_{y'}$, the fifth inequality holds since (3.5.24) implies that there is at most one $x' \in V(H)$ with $\phi_*(x') = v$, and the sixth inequality holds since $|J^{\vec{s}}_y| \leq |N_{H^{\vec{s}}_y}(x^{\vec{s}}_y)| \leq \Delta$.)

Consider any $(\vec{s}, y) \in V(\Lambda)$. Then similarly as above we have

$$d_\Lambda((\vec{s}, y)) \leq \sum_{v \in \phi_*(J^{\vec{s}}_y)} \sum_{H \in \mathcal{H}} |\{x \in N_H(x') : v = \phi_*(x'), x' \in V(H)\}| \leq \Delta^2 |\mathcal{H}| \overset{(3.5.2)}{\leq} \alpha^{1/2} n'.$$

This shows that

$$\Delta(\Lambda) \leq \alpha^{1/2} n' < d^{5\Delta} n'/2. \quad (3.5.37)$$

We can now apply the blow-up lemma for approximate decompositions (Theorem 3.3.15) with the following objects and parameters.

object/parameter	$G[V]$	V_i	$\widehat{X}_i$	$H_{t,w'}$	$\mathbf{S}= [T] \times [w]$	q	$T^{-1/2}$	Q
playing the role of	G	V_i	X_i	H_i	$[s]$	q	ε	R
object/parameter	r	$L_i^{t,w'} \cup M_i^{t,w'}$	$A_y^{t,w'}$	α	$d^{5\Delta}$	ν	Λ	n'
playing the role of	r	W_i^j	A_w^j	d	d_0	α	Λ	n

Indeed, $(A3)_{3.5.1}$ implies that $(A1)_{3.3.15}$ holds, and $(A2)_{3.3.15}$ holds by the definition of $H_{t,w'}$. Claim 3.5.5 and (3.5.35) imply that $(A3)_{3.3.15}$ holds, and (3.5.35), (3.5.36) and (3.5.37) imply that $(A4)_{3.3.15}$ holds. Moreover, (3.5.25) implies that the upper bound on s in the assumption of Theorem 3.3.15 holds.

Thus by Theorem 3.3.15 we obtain a function ϕ^* that packs $\{H_{\vec{s}} : \vec{s} \in \mathbf{S}\}$ into $G[V]$ and satisfies the following, where $\phi^*_{\vec{s}}$ denotes the restriction of ϕ^* to $H_{\vec{s}}$.

$(\Phi^*1)_{3.5.1}$ for each $\vec{s} \in \mathbf{S}$ and $y \in \bigcup_{i\in[r]} L_i^{\vec{s}} \cup M_i^{\vec{s}}$, we have $\phi^*_{\vec{s}}(y) \in A_y^{\vec{s}}$,

$(\Phi^*2)_{3.5.1}$ for any $(\vec{s}, y)(\vec{t}, y') \in E(\Lambda)$, we have that $\phi^*_{\vec{s}}(y) \neq \phi^*_{\vec{t}}(y')$.

We let

$$\phi := \phi^*(\bigcup_{\vec{s}\in\mathbf{S}} \Phi_s) \cup \phi_*.$$

Recall from Step 3 and (3.5.23) that $\phi_* = \phi' \cup \phi''$, and that ϕ' packs $\{H[Y^H \cup Z^H \cup A^H] : H \in \mathcal{H}\}$ into $G[U] \cup F'$. Since each $\Phi_{\vec{s}}$ is a packing of $\{H[X^H \setminus W^H] : H \in \mathcal{H}_{\vec{s}}\}$ into $H_{\vec{s}}$ and ϕ^* is a packing of $\{H_{\vec{s}} : \vec{s} \in \mathbf{S}\}$ into $G[V]$, we know that ϕ packs $\{H[X^H \setminus W^H] : H \in \mathcal{H}\}$ into $G[V]$. Moreover, $(\Phi^*1)_{3.5.1}$, $(\Phi^*2)_{3.5.1}$ with the definitions of $A_y^{\vec{s}}$ and Λ imply that ϕ packs $\{H[X^H \setminus W^H, W^{H,F}] : H \in \mathcal{H}\}$ into F, and ϕ packs $\{H[X^H \setminus W^H, W^{H,U'}] : H \in \mathcal{H}\}$ into $G[V, U_0]$, and ϕ packs $\{H[X^H \setminus W^H, W^{H,D} \cup Z^H] : H \in \mathcal{H}\}$ into $G[V, U]$. Thus, we have the following.

$$\phi(\bigcup_{H\in\mathcal{H}} E_H(Y^H \cup Z^H \cup A^H)) \subseteq E_G(U) \cup E(F'), \qquad \phi(\bigcup_{H\in\mathcal{H}} E_H(X^H \setminus W^H)) \subseteq E_G(V),$$
$$\phi(\bigcup_{H\in\mathcal{H}} E_H(X^H \setminus W^H, W^{H,F})) \subseteq E_F(V, U), \qquad \phi(\bigcup_{H\in\mathcal{H}} E_H(X^H \setminus W^H, W^{H,U'})) \subseteq E_G(V, U_0),$$

$$\phi(\bigcup_{H\in\mathcal{H}} E_H(X^H \backslash W^H, W^{H,D} \cup Z^H)) \subseteq E_G(V,U). \tag{3.5.38}$$

Also, it is obvious that the restriction of ϕ to $V(H)$ is injective for each $H \in \mathcal{H}$. As $G[U] \cup F', G[V], F, G[V,U_0]$ and $G[V,U]$ are pairwise edge-disjoint, we conclude that ϕ packs $\mathcal{H}$ into $G \cup F \cup F'$. Moreover, by (3.5.2) we have $\Delta(\phi(\mathcal{H})) \leq \Delta|\mathcal{H}| \leq 4k\Delta\alpha n/r$, thus $(\mathrm{B1})_{3.5.1}$ holds. By (3.5.38) and $(\Phi'4)_{3.5.1}$, for $u \in U$, we have

$$d_{\phi(\mathcal{H})\cap G}(u) \leq \Delta|\{H \in \mathcal{H} : u \in \phi_*(Y^H \cup Z^H \cup W^{H,D})\}| \overset{(\Phi_*4)_{3.5.1}}{\leq} \frac{2\Delta\varepsilon^{1/8}n}{r}.$$

Thus $(\mathrm{B2})_{3.5.1}$ holds.

Finally, for $i \in [r]$, by $(\mathrm{X3})_{3.5.1}$, $(\mathrm{X6})_{3.5.1}$, (3.5.38) we have

$$\begin{aligned} e_{\phi(\mathcal{H})\cap G}(V_i, U \cup U_0) &\leq \sum_{H\in\mathcal{H}} \Delta\Big(|Z^H| + \sum_{j\in N_Q(i)} |W_j^{H,U'}| + \sum_{j\in N_Q(i)} |W_j^{H,D}|\Big) \\ &\overset{(3.5.2),(\mathrm{X6})_{3.5.1},\ (3.5.6),(3.5.19)}{\leq} \frac{2k\Delta\alpha n}{r}\Big(4\Delta^{3k^3}\eta^{0.9}n + 2(k-1)\varepsilon^{3/4}n/r + (k-1)r\eta^{1/7}n\Big) \\ &\leq \frac{\varepsilon^{1/2}n^2}{r^2}, \end{aligned}$$

which shows that $(\mathrm{B3})_{3.5.1}$ holds. □

3.6 Proof of Theorem 3.1.2

The proof of Theorem 3.1.2 proceeds in three steps. In the first step we will apply the results of Section 3 to construct suitable edge-disjoint subgraphs $G_{t,s}, G_t^*, F_{t,s}$ and F_t' of G, where $G_{t,s}$ is a K_k-factor blow-up spanning almost all vertices while $G_t^*, F_{t,s}$ and F_t' are comparatively sparse. In the (straightforward) second step, we simply partition $\mathcal{H}$ into collections $\mathcal{H}_{t,s}$ such that the $e(\mathcal{H}_{t,s})$ are approximately equal to each other. Finally, in the third step we will pack each $\mathcal{H}_{t,s}$ into $G_{t,s} \cup G_t^* \cup F_{t,s} \cup F_t'$ via Lemma 3.5.1.

Proof of Theorem 3.1.2. Let $\sigma := \delta - \max\{1/2, \delta_k^{\mathrm{reg}}\} > 0$. By (3.3.6), we have $\delta \geq 1 - 1/k + \sigma$ for any $k \geq 2$. Without loss of generality, we may assume that $\nu < \sigma/2$. For given $\nu, \sigma > 0$

and Δ, $k \in \mathbb{N}\backslash\{1\}$, we choose constants $n_0, \xi, \eta, M, M', \varepsilon, T, q, d$ such that $q \mid T$ and

$$0 < 1/n_0 \ll \eta \ll 1/M \ll 1/M' \ll \varepsilon \ll 1/T \ll 1/q \ll \xi \ll d \ll \nu, \sigma, 1/\Delta, 1/k \leq 1/2. \tag{3.6.1}$$

Suppose $n \geq n_0$ and let G be an n-vertex graph satisfying condition (i) of Theorem 3.1.2. Furthermore, suppose $\mathcal{H}$ is a collection of (k, η)-chromatic η^2-separable graphs satisfying conditions (ii) and (iii) of Theorem 3.1.2. We will show that $\mathcal{H}$ packs into G. Note that we assume $\mathcal{H}$ to consist of η^2-separable graphs here (instead of η-separable graphs). This is more convenient for our purposes, but still implies Theorem 1.2.

Step 1. Decomposing G into host graphs. In this step, we apply Szemerédi's regularity lemma to G and then apply Lemma 3.3.16 to obtain a partition of $V(G)\backslash V_0$ into T reservoir sets Res_t, where V_0 is the exceptional set obtained from Szemerédi's regularity lemma. We use Lemma 3.3.13 to obtain an approximate decomposition of the reduced multi-graph R'_{multi} of G into almost K_k-factors and partition these factors into T collections. Each such almost K_k-factor Q gives us an ε-regular Q-blow-up $G_{t,s}$ in G, and we modify it into a super-regular Q-blow-up. We also put aside several sparse 'connection graphs' $F_{t,s}$ and F'_t, which will be used to link vertices in the reservoir and exceptional set with vertices in the rest of the graph. These connection graphs will play the roles of F and F' in Lemma 3.5.1. We also put aside a further sparse connection graph G^*_t which provides additional connections within $V(G) \setminus V_0$.

We apply Szemerédi's regularity lemma (Lemma 3.3.5) with (ε^2, d) playing the role of (ε, d) to obtain a partition $V'_0, \ldots, V'_{r'}$ of $V(G)$ and a spanning subgraph $G' \subseteq G$ such that

(R1) $M' \leq r' \leq M$,

(R2) $|V'_0| \leq \varepsilon^2 n$,

(R3) $|V'_i| = |V'_j| = (1 \pm \varepsilon^2)n/r'$ for all $i, j \in [r']$,

(R4) for all $v \in V(G)$ we have $d_{G'}(v) > d_G(v) - 2dn$, (R5) $e(G'[V_i']) = 0$ for all $i \in [r']$,

(R6) for any i, j with $1 \le i \le j \le r'$, the graph $G'[V_i', V_j']$ is either empty or $(\varepsilon^2, d_{i,j})$-regular for some $d_{i,j} \in [d, 1]$.

Let R' be the graph with

$$V(R') = [r'] \qquad \text{and} \qquad E(R') := \{ij : e_{G'}(V_i', V_j') > 0\}.$$

Note that for $i, j \in [r']$, $ij \in E(R')$ if and only if $G'[V_i', V_j']$ is $(\varepsilon^2, d_{i,j})$-regular with $d_{i,j} \ge d$. Now, we let R'_{multi} be a multi-graph with $V(R'_{\text{multi}}) = [r']$ and with exactly

$$q_{i,j} := \lfloor (1 - 6d) d_{i,j} q \rfloor \tag{3.6.2}$$

edges between i and j for each $ij \in E(R')$. Note that R'_{multi} has edge-multiplicity at most q. For each $i \in [r']$, we have

$$\begin{aligned} d_{R'_{\text{multi}}}(i) &= \sum_{j \in N_{R'}(i)} \lfloor (1-6d) q (\frac{e_{G'}(V_i', V_j')}{|V_i'||V_j'|} \pm \varepsilon^2) \rfloor \overset{\text{(R3),(R5)}}{=} \frac{\sum_{v \in V_i'} (1-6d) q d_{G', V(G) \setminus V_0}(v)}{|V_i'|^2} \pm \varepsilon^2 q r' \pm r' \\ &\overset{\text{(R2),(R4)}}{=} \frac{q}{|V_i'|^2} \sum_{v \in V_i'} (d_G(v) \pm 10dn) \pm 2r' \overset{(i)}{=} \frac{(\delta \pm 11d) q n}{|V_i'|} \pm 2r' \overset{\text{(R3)}}{=} (\delta \pm d^{3/4}) q r'. \end{aligned} \tag{3.6.3}$$

We apply Lemma 3.3.13 with $R'_{\text{multi}}, r', \varepsilon^2, k, \sigma, d^{3/4}, \nu/5, T$ and q playing the roles of G,n,ε, k, σ, ξ, ν, T and q, respectively. Then, by permuting indices in $[r']$ if necessary, we obtain $R_{\text{multi}} \subseteq R'_{\text{multi}}$ and a collection $\mathcal{Q} := \{Q_{1,1}, \dots, Q_{1,\kappa/T}, Q_{2,1}, \dots, Q_{T,\kappa/T}\}$ of edge-disjoint subgraphs of R_{multi} such that the following hold.

(Q1) $R_{\text{multi}} = R'_{\text{multi}}[[r]]$ with $(1 - \varepsilon^2) r' \le r \le r'$, and $k \mid r$,

(Q2) $\kappa = \frac{(\delta - \nu/5 \pm \varepsilon^2) q r'}{k-1} = \frac{(\delta - \nu/5 \pm \varepsilon) q r}{k-1}$ and $T \mid \kappa$,

(Q3) for each $(t, s) \in [T] \times [\kappa/T]$, $Q_{t,s}$ is a vertex-disjoint union of at least $(1 - \varepsilon) r/k$ copies of K_k,

(Q4) for each $i \in [r]$, we have $|\{(t, s) \in [T] \times [\kappa/T] : i \in V(Q_{t,s})\}| \ge \kappa - \varepsilon r$, (Q5) for all $t \in [T]$ and $i, j \in [r]$, we have $|\{s \in [\kappa/T] : j \in N_{Q_{t,s}}(i)\}| \leq 1$.

For each $t \in [T]$, let $\mathcal{Q}_t := \{Q_{t,1}, \ldots, Q_{t,\kappa/T}\}$. We define $R := R'[[r]]$ to be the induced subgraph of R' on $[r]$. Note that each $Q_{t,s} \in \mathcal{Q}$ can be viewed as a subgraph of R. Moreover, for fixed $t \in [T]$, (Q5) implies that the graphs $Q_{t,1}, \ldots, Q_{t,\kappa/T}$ are pairwise edge-disjoint when viewed as subgraphs of R. Also, we have

$$\delta(R) \geq q^{-1}\delta(R'_{\text{multi}}) - (r' - r) \overset{(3.6.3),(Q1)}{\geq} (\delta - d^{1/2})r. \tag{3.6.4}$$

We need to modify the sets V_i' later to ensure that we obtain appropriate super-regular $Q_{t,s}$-blow-ups. For this, we need to move some 'bad' vertices in V_i' into V_0'. For each $i \in [r]$ and each $j \in N_R(i)$, we define

$$U_i(j) := \{v \in V_i' : d_{G',V_j'}(v) \neq (d_{i,j} \pm \varepsilon^2)|V_j'|\} \text{ and } U_i' := \{v \in V_i' : |\{j : v \in U_i(j)\}| > \varepsilon r\}. \tag{3.6.5}$$

By Proposition 3.3.4 and (R6), for any $i \in [r]$ and $j \in N_R(i)$ we have

$$|U_i(j)| \leq 5\varepsilon^2 n/r \qquad \text{and} \qquad |U_i'| \leq (\varepsilon r)^{-1} \sum_{j \in N_R(i)} |U_i(j)| \leq 5\varepsilon n/r. \tag{3.6.6}$$

For each $i \in [r]$, we let $V_i := V_i' \setminus U_i'$ and $V_0 := V_0' \cup \bigcup_{i=1}^{r} U_i' \cup \bigcup_{i=r+1}^{r'} V_i'$.

By (R2) and (R3), for each $i \in [r]$, we have

$$(1 - 6\varepsilon)n/r \leq |V_i| \leq n/r \qquad \text{and} \qquad |V_0| \leq 6\varepsilon n. \tag{3.6.7}$$

We apply Lemma 3.3.16 with $G', V(G)\setminus V_0, \{V_i\}_{i=1}^r$ and T playing the roles of $G, V, \{V_i\}_{i=1}^r$ and t to obtain a partition $\{Res_1, \ldots, Res_T\}$ of $V(G)\setminus V_0$ satisfying the following, where we define $V_i^t := V_i \cap Res_t$.

(Res1) For all $t \in [T]$ and $v \in V(G)$, we have $d_{G',V_i^t}(v) = \frac{1}{T} d_{G',V_i}(v) \pm n^{2/3}$,

(Res2) for all $t \in [T]$ and $i \in [r]$, we have $|V_i^t| = (\frac{1}{T} \pm \varepsilon^2)|V_i| \overset{(3.6.7)}{=} \frac{(1 \pm 7\varepsilon)n}{Tr}$,

(Res3) for all $t \in [T]$, we have $|Res_t| \in \{\lfloor \frac{n - |V_0|}{T} \rfloor, \lfloor \frac{n - |V_0|}{T} \rfloor + 1\}$.

Next, we partition the edges in $G' \setminus V_0$ into $L_1, \dots, L_7$ which will be the building blocks for the graphs G, F and F' in Lemma 3.5.1. Let $p_1 := 1 - 6d$ and $p_j := d$ for $2 \leq j \leq 7$. Apply Lemma 3.3.17 with $G' \setminus V_0$, $\{V_i^t : i \in [r], t \in [T]\}$, $\{(V_i, V_j) : ij \in E(R)\}$ and 7 playing the roles of $G, \mathcal{U}, \mathcal{U}'$ and s. Then we obtain a decomposition $L_1, \dots, L_7$ of $G' \setminus V_0$ satisfying the following for all $t \in [T]$, $i \in [r]$, $\ell \in [7]$ and $v \in V(G) \setminus V_0$:

(L1) $d_{L_\ell, V_i^t}(v) = p_\ell d_{G', V_i^t}(v) \pm n^{2/3}$,

(L2) for each $ij \in E(R)$, we have that $L_\ell[V_i, V_j]$ is $(4\varepsilon^2, d_{i,j}p_\ell)$-regular.

Let $G'' := L_1$. For each $t \in [T]$, let G_t^*, F_t and F_t^* be the graphs on vertex set $V(G) \setminus V_0$ with

$$E(G_t^*) := \bigcup_{t'=1}^{t-1} E(L_2[Res_t, Res_{t'}]) \cup \bigcup_{t'=t+1}^{T} E(L_3[Res_t, Res_{t'}]) \cup L_2[Res_t], \tag{3.6.8}$$
$$E(F_t) := \bigcup_{t'=1}^{t-1} E(L_4[Res_t, Res_{t'}]) \cup \bigcup_{t'=t+1}^{T} E(L_5[Res_t, Res_{t'}]),$$
$$E(F_t^*) := \bigcup_{t'=1}^{t-1} E(L_6[Res_t, Res_{t'}]) \cup \bigcup_{t'=t+1}^{T} E(L_7[Res_t, Res_{t'}]).$$

For each $t \in [T]$, we let $F_{t,1}, \dots, F_{t,\kappa/T}$ be subgraphs of F_t such that for all $s \in [\kappa/T]$

$$F_{t,s} := \bigcup_{i \in V(Q_{t,s})} \bigcup_{j \in N_{Q_{t,s}}(i)} F_t[V_i^t, V_j \setminus Res_t]. \tag{3.6.9}$$

Note that (Q5) implies that for $s \neq s' \in [\kappa/T]$, the graphs $F_{t,s}$ and $F_{t,s'}$ are edge-disjoint. Thus $G'', G_1^*, \dots, G_T^*, F_{1,1}, \dots, F_{T,\kappa/T}, F_1^*, \dots, F_T^*$ form edge-disjoint subgraphs of $G' \setminus V_0$. The edges in G_t^* will be used to satisfy condition $(A4)_{3.5.1}$ when applying Lemma 3.5.1. The graphs $F_{t,s}$ will play the role of F in Lemma 3.5.1. The graphs F_t^* will be used in the construction of the graph F_t', which will play the role of F' in Lemma 3.5.1.

We will now further partition the edges in $G'' = L_1$. Note that for each $ij \in E(R)$, by (3.6.2) we have $q_{i,j} = \lfloor d_{i,j} p_1 q \rfloor$. To further partition G'', we apply Lemma 3.3.17 for each $ij \in E(R)$ with the following objects and parameters.

object/parameter	$G''[V_i,V_j]$	$\{V_i^t, V_j^t : t \in [T]\}$	$\{(V_i,V_j)\}$	$q_{i,j}+1$	$1/(d_{i,j}p_1q)$	$1-q_{i,j}/(d_{i,j}p_1q)$
playing the role of	G	$\mathcal{U}$	$\mathcal{U}'$	s	$p_i : i < s$	p_s

Then by (L2), for each $ij \in E(R)$, we obtain edge-disjoint subgraphs $E^1_{i,j}, \dots, E^{q_{i,j}+1}_{i,j}$ of $G''[V_i,V_j]$ satisfying the following for all $t \in [T]$ and $\ell \in [q_{i,j}]$:

(E1) for each $v \in V_i$, we have $d_{E^\ell_{i,j},V_j^t}(v) = \frac{1}{d_{i,j}p_1q} d_{G'',V_j^t}(v) \pm n^{2/3}$,

(E2) $E^\ell_{i,j}$ is $(8\varepsilon^2, 1/q)$-regular.

Recall that we have chosen a collection $\mathcal{Q} = \{Q_{1,\kappa/T}, \dots, Q_{T,\kappa/T}\}$ of edge-disjoint subgraphs of R_{multi} satisfying (Q1)–(Q5). Let $\psi : E(R_{\text{multi}}) \to \mathbb{N}$ be a function such that

$$\psi(E_{R_{\text{multi}}}(i,j)) = [q_{i,j}].$$

For all $ij \in E(R')$, there are exactly $q_{i,j}$ edges between i and j in R_{multi}, so such a function ψ exists. Now, for all $t \in [T]$, $s \in [\kappa/T]$, we let

$$G_{t,s} := \bigcup_{ij \in E(Q_{t,s})} E^{\psi(ij)}_{i,j}. \tag{3.6.10}$$

Since $\mathcal{Q}$ is a collection of edge-disjoint subgraphs of R_{multi} and $E^1_{i,j}, \dots, E^{q_{i,j}+1}_{i,j}$ are edge-disjoint subgraphs of G'', the graphs $G_{1,1}, \dots, G_{T,\kappa/T}$ form edge-disjoint subgraphs of G''.

We would like to use $G_{t,s} \setminus Res_t$ and Res_t to play the roles of $G[\bigcup_{i\in[r]} V_i]$ and U in Lemma 3.5.1, respectively. However, $E^\ell_{i,j} \setminus Res_t$ is not necessarily super-regular and the sizes of $V_i \setminus Res_t$ are not necessarily the same for all $i \in [r]$. To ensure this, we will now choose an appropriate subset $V_i^{t,s}$ of V_i which can play the role of V_i in Lemma 3.5.1.

For all $t \in [T]$, $i \in [r]$ and $s \in [\kappa/T]$, let

$$V_i(t,s) := V_i \setminus (Res_t \cup \bigcup_{j \in N_{Q_{t,s}}(i)} U_i(j)) \quad \text{and} \quad m := \frac{(T-1)n}{Tr} - \frac{10\varepsilon n}{r}. \tag{3.6.11}$$

Then by (3.6.6), (3.6.7) and (Res2), we have

$$0 \le |V_i(t,s)| - m \le 15\varepsilon n/r. \tag{3.6.12}$$

For all $t \in [T]$ and $i \in [r]$, we apply Lemma 3.3.18 with the following objects and parameters.

object/parameter	κ/T	1	$s \in [\kappa/T]$	$V_i \setminus Res_t$	$V_i(t,s)$	20ε	$\lvert V_i(t,s)\rvert - m$	d
playing the role of	s	r	$i \in [s]$	A	$A_{i,1}$	ε	$m_{i,1}$	$1/2$

Then we obtain sets $W_i(t,1),\dots,W_i(t,\kappa/T)$ such that $W_i(t,s) \subseteq V_i(t,s)$ with $|V_i(t,s)\setminus W_i(t,s)| = m$ and for any $v \in V_i \setminus Res_t$, we have

$$|\{s \in [\kappa/T] : v \in W_i(t,s)\}| \le 10\varepsilon^{1/2}\kappa/T. \tag{3.6.13}$$

For all $t \in [T]$, $s \in [\kappa/T]$ and $i \in V(Q_{t,s})$, let $V_i^{t,s} := V_i(t,s) \setminus W_i(t,s)$. Let

$$V_0^{t,s} := V_0 \cup \bigcup_{i\in[r]} \bigcup_{j\in N_{Q_{t,s}}(i)} (U_i(j) \setminus Res_t) \cup \bigcup_{i\in[r]} W_i(t,s) \cup \bigcup_{i\in[r]\setminus V(Q_{t,s})} (V_i \setminus Res_t).$$

Then the sets $V_0^{t,s}, \{V_i^{t,s} : i \in V(Q_{t,s})\}, Res_t$ form a partition of $V(G)$, and for each $i \in V(Q_{t,s})$

$$|V_i^{t,s}| = m := \frac{(T-1)n}{Tr} - \frac{10\varepsilon n}{r}, \text{ and} \tag{3.6.14}$$

$$\begin{aligned} |V_0^{t,s}| &\overset{(3.6.6),(3.6.7),(3.6.12)}{\le} 6\varepsilon n + (k-1)r(5\varepsilon^2 n/r) + 15\varepsilon n + (r - |V(Q_{t,s})|)n/r \\ &\overset{(Q3)}{\le} 25\varepsilon n. \end{aligned} \tag{3.6.15}$$

We now further modify V_i^t into $U_i^{t,s}$ which can play the role of U_i in Lemma 3.5.1. For all $(t,s) \in [T]\times[\kappa/T]$ and $i \in V(Q_{t,s})$, we define

$$U_i^{t,s} := V_i^t \setminus \bigcup_{j\in N_{Q_{t,s}}(i)} U_i(j) \text{ and } U_0^{t,s} := \bigcup_{i\in[r]\setminus V(Q_{t,s})} V_i^t \cup \bigcup_{i\in V(Q_{t,s})} \bigcup_{j\in N_{Q_{t,s}}(i)} U_i(j).$$

Note that for each $(t,s) \in [T]\times[\kappa/T]$, the sets $\{U_0^{t,s}\} \cup \{U_i^{t,s} : i \in V(Q_{t,s})\}$ form a partition of Res_t. By (3.6.6), for all $(t,s) \in [T]\times[\kappa/T]$ and $i \in V(Q_{t,s})$, we have

$$|U_i^{t,s}| = |V_i^t| \pm 5k\varepsilon^2 n/r \overset{(Res2)}{=} \frac{(1\pm 8\varepsilon)n}{Tr} \text{ and } |U_0^{t,s}| \overset{(3.6.6)}{\le} \sum_{i\in[r]\setminus V(Q_{t,s})} |V_i^t| + 5k\varepsilon^2 n \overset{(Q3)}{\le} 2\varepsilon n. \tag{3.6.16}$$

Note that for all $(t,s) \in [T] \times [\kappa/T]$ and $i \in V(Q_{t,s})$, we have $U_i^{t,s}, V_i^{t,s} \subseteq V_i$. Thus Proposition 3.3.2 with (3.6.14), (3.6.16), (L2) and the definition of p_ℓ implies that for all $(t,s) \in [T] \times [\kappa/T]$, $ij \in E(R[V(Q_{t,s})])$ and $i'j' \in E(Q_{t,s})$, we have

$$G_t^*[U_i^{t,s}, U_j^{t,s}], G_t^*[V_i^{t,s}, U_j^{t,s}] \text{ and } F_{t,s}[V_{i'}^{t,s}, U_{j'}^{t,s}] \text{ are } (\varepsilon, (d^2))^+\text{-regular.} \tag{3.6.17}$$

Moreover, for all $(t,s) \in [T] \times [\kappa/T]$, $ij \in E(Q_{t,s})$ and $u \in U_i^{t,s}$, we have

$$\begin{aligned} d_{F_{t,s}, V_j^{t,s}}(u) &\overset{(3.6.9),(3.6.14),(\text{Res2})}{\geq} d_{F_t[V_i^t, V_j \setminus Res_t]}(u) - n/(Tr) \overset{(\text{L1}),(\text{Res1})}{\geq} d \cdot d_{G', V_j}(u) - 3n/(Tr) \\ &\overset{(3.6.5),(3.6.6)}{\geq} d \cdot (d_{i,j} - \varepsilon^2)|V_j| - 4n/(Tr) \overset{(\text{Res2})}{\geq} (2d^2/3)|V_j \setminus Res_t|. \end{aligned} \tag{3.6.18}$$

We obtain the third inequality from the definition of $U_i^{t,s}$ and the fact that $ij \in E(Q_{t,s})$.

Claim 3.6.1. *For all $t \in [T], s \in [\kappa/T]$ and $ij \in E(Q_{t,s})$, the graph $G_{t,s}[V_i^{t,s}, V_j^{t,s}]$ is $(\varepsilon^{1/2}, 1/q)$-super-regular.*

Proof. Let $\ell \in [q_{i,j}]$ be such that $G_{t,s}[V_i, V_j] = E_{i,j}^\ell$. Such an ℓ exists by the definition of $G_{t,s}$ and the assumption that $ij \in E(Q_{t,s})$. Note that for $i' \in \{i,j\}$ we have $V_{i'}^{t,s} \subseteq V_{i'}$ with $|V_{i'}^{t,s}| = m > \frac{1}{2}|V_{i'}|$ by (3.6.14). Thus Proposition 3.3.2 with (E2) implies that $G_{t,s}[V_i^{t,s}, V_j^{t,s}] = E_{i,j}^\ell[V_i^{t,s}, V_j^{t,s}]$ is $(16\varepsilon^2, 1/q)$-regular.

Consider $v \in V_i^{t,s}$. By the definition of $V_i^{t,s}$, we have $v \notin U_i(j)$. Thus

$$\begin{aligned} d_{G_{t,s}, V_j^{t,s}}(v) &\overset{(3.6.6),(3.6.12)}{=} d_{E_{i,j}^\ell, V_j \setminus Res_t}(v) \pm \frac{16\varepsilon n}{r} = \sum_{t' \in [T]\setminus\{t\}} d_{E_{i,j}^\ell, V_j^{t'}}(v) \pm \frac{16\varepsilon n}{r} \\ &\overset{(\text{E1})}{=} \sum_{t' \in [T]\setminus\{t\}} \frac{1}{d_{i,j} p_1 q} d_{G'', V_j^{t'}}(v) \pm \frac{17\varepsilon n}{r} \overset{(\text{L1})}{=} \sum_{t' \in [T]\setminus\{t\}} \frac{1}{d_{i,j} q} d_{G', V_j^{t'}}(v) \pm \frac{18\varepsilon n}{r} \\ &\overset{(\text{Res1})}{=} \frac{(T-1)}{d_{i,j} qT} d_{G', V_j}(v) \pm \frac{19\varepsilon n}{r} \overset{(3.6.5)}{=} \frac{(T-1)}{d_{i,j} qT}((d_{i,j} \pm \varepsilon^2)|V_j'| \pm |U_j'|) \pm \frac{19\varepsilon n}{r} \\ &\overset{(3.6.6)}{=} \frac{(T-1)n}{qTr} \pm \frac{30\varepsilon n}{r} \overset{(3.6.14)}{=} (\frac{1}{q} \pm \varepsilon^{1/2})|V_j^{t,s}|. \end{aligned}$$

Similarly, for $v \in V_j^{t,s}$, we have $d_{G_{t,s}, V_i^{t,s}}(v) = (\frac{1}{q} \pm \varepsilon^{1/2})|V_i^{t,s}|$. Thus $G_{t,s}[V_i^{t,s}, V_j^{t,s}]$ is $(\varepsilon^{1/2}, 1/q)$-super-regular. This proves the claim. □

For all $t \in [T]$, $v \in Res_t$ and $s \in [\kappa/T]$, we know that

$$d_{G_t^*, V_i^{t,s}}(v) = d_{G_t^*, V_i}(v) \pm |V_i \setminus V_i^{t,s}| \overset{(L1)}{=} \sum_{\ell \in [T]} (d \cdot d_{G', V_i^\ell}(v) \pm n^{2/3}) \pm |V_i \setminus V_i^{t,s}|$$
$$\overset{(Res1),(3.6.14)}{=} d \cdot d_{G', V_i}(v) \pm 2n/(Tr).$$

This implies that

$$|\{i \in V(Q_{t,s}) : d_{G_t^*, V_i^{t,s}}(v) \geq d^2 m/2\}| \geq |\{i \in V(Q_{t,s}) : d_{G', V_i}(v) \geq d|V_i|\}|$$
$$\geq \frac{d_{G'}(v) - |V_0| - dn}{\max_{i \in [r]} |V_i|} - |[r] \setminus V(Q_{t,s})| \overset{(3.6.7),(Q3)}{\geq} (1 - 1/k + \sigma/2)r. \tag{3.6.19}$$

We obtain the final inequality since $\delta(G') \geq (\delta - \xi - 2d)n \geq (1 - 1/k + 3\sigma/4)n$ by (i) and (R4). This together with (3.6.17) and Claim 3.6.1 will ensure that $G_{t,s} \cup G_t^*$ can play the role of G in Lemma 3.5.1, and (3.6.18) shows that $F_{t,s}$ can play the role of F in Lemma 3.5.1.

The remaining part of this step is to construct a graph which can play the role of F' in Lemma 3.5.1. F' needs to contain suitable stars centred at v whenever $v \in V_0^{t,s}$. (For each t, the number of stars we will need for v in order to deal with all $s \in [\kappa/T]$ is bounded from above by (3.6.23).) For all $t \in [T]$, $s \in [\kappa/T]$, $v \in V(G)$ and $u \in Res_t$, let

$$I_t(v) := \{s' \in [\kappa/T] : v \in V_0^{t,s'}\} \quad \text{and} \quad i_t^s(v) := |I_t(v) \cap [s]|,$$
$$J_t(u) := \{s' \in [\kappa/T] : u \in U_0^{t,s'}\} \quad \text{and} \quad j_t^s(u) := |J_t(u) \cap [s]|. \tag{3.6.20}$$

Note that if $v \in V_0$, then $I_t(v) = [\kappa/T]$. If $v \in V_i \backslash Res_t$ for some $i \in [r]$, then $s \in I_t(v)$ means $v \in W_i(t,s) \cup \bigcup_{j \in N_{Q_{t,s}}(i)} U_i(j) \cup \bigcup_{i' \in [r] \setminus V(Q_{t,s})} V_{i'}$. Together with the fact that $U_i' \subseteq V_0$ and so $v \notin U_i'$, this implies

$$|I_t(v)| \overset{(Q5)}{\leq} |\{s \in [\kappa/T] : v \in W_i(t,s)\}| + |\{j \in [r] : v \in U_i(j)\}| + |\{s \in [\kappa/T] : i \notin V(Q_{t,s})\}|$$
$$\overset{(3.6.5),(3.6.13),(Q4)}{\leq} 10\varepsilon^{1/2}\kappa/T + \varepsilon r + \varepsilon r \overset{(Q2)}{\leq} 20\varepsilon^{1/2} r. \tag{3.6.21}$$

Similarly, for $u \in V_i^t$, we have

$$|J_t(u)| \leq |\{j \in [r] : u \in U_i(j)\}| + |\{s \in [\kappa/T] : i \notin V(Q_{t,s})\}| \overset{(3.6.5),(Q4)}{\leq} \varepsilon r + \varepsilon r \leq 2\varepsilon r. \tag{3.6.22}$$

For each $v \in V(G) \setminus Res_t$, let

$$\kappa_v := \begin{cases} (1+d)\kappa & \text{if } v \in V_0, \\ \lceil r/(2k) \rceil & \text{if } v \notin V_0. \end{cases} \tag{3.6.23}$$

κ_v is the overall number of stars centred at v that we will construct for given t. Note that for all $t \in [T]$ and $s \in [\kappa/T]$, no edge of $E(G'[V_0, Res_t])$ belongs to any of the graphs $G_{t,s}, G_t^*, F_t, F_t^*$. Now for each $t \in [T]$, we use these edges and edges in F_t^* to construct stars $F_t'(v,s)$ centred at v, and subsets $C_{v,s}^t$, $C_{v,s}^{*,t}$ of $[r]$ for all $v \in V(G) \setminus Res_t$ and $s \in [\kappa_v]$, in such a way that the following hold for all $t \in [T]$ and $v \in V(G) \setminus Res_t$.

(F′1) For each $s \in [\kappa_v]$, we have $C_{v,s}^t \subseteq C_{v,s}^{*,t}$, $|C_{v,s}^t| = k-1$, $|C_{v,s}^{*,t}| = k$ and $R[C_{v,s}^{*,t}] \simeq K_k$,

(F′2) for each $i \in [r]$, we have $|\{s \in [\kappa_v] : i \in C_{v,s}^{*,t}\}| \leq (k+1)q$,

(F′3) for each $s \in [\kappa_v]$, if $i \in C_{v,s}^t$, then $d_{F_t'(v,s),V_i^t}(v) \geq \frac{|V_i^t|}{q}$.

Claim 3.6.2. *For all $t \in [T]$, $v \in V(G) \setminus Res_t$ and $s \in [\kappa_v]$, there exist edge-disjoint stars $F_t'(v,s) \subseteq G'[V_0, Res_t] \cup F_t^*$ centred at v, and subsets $C_{v,s}^t$, $C_{v,s}^{*,t}$ of $[r]$ which satisfy (F′1)–(F′3).*

When applying Lemma 3.5.1 in Step 3 to pack $\mathcal{H}_{t,s}$, we will only make use of those stars $F_t'(v,s)$ with $v \in V_0^{t,s}$, but it is slightly more convenient to define them for all $v \in V(G) \setminus Res_t$.

Proof. First, consider $t \in [T]$ and $v \in V_0$. Then we have

$$d_{G',Res_t}(v) = \sum_{i \in [r]} d_{G',V_i^t}(v) \overset{\text{(Res1)}}{=} \frac{1}{T} \sum_{i \in [r]} d_{G',V_i}(v) \pm rn^{2/3} \overset{(i),\text{(R4)},\text{(Res3)},(3.6.7)}{=} (\delta \pm 3d)|Res_t|. \tag{3.6.24}$$

For all $v \in V_0$, $t \in [T]$ and $i \in [r]$, let $q_{v,i}^t := \lfloor \frac{q \cdot d_{G',V_i^t}(v)}{|V_i^t|} \rfloor$. Consider edge-disjoint subsets $E_{v,i}^t(1), \ldots, E_{v,i}^t(q_{v,i}^t)$ of $E_{G'}(\{v\}, V_i^t)$ such that $|E_{v,i}^t(q')| = \frac{1}{q}|V_i^t|$ for each $q' \in [q_{v,i}^t]$. Let R_v^t be an auxiliary graph such that

$$V(R_v^t) := \{(i,q') : i \in [r], q' \in [q_{v,i}^t]\} \quad \text{and} \quad E(R_v^t) := \{(i,q')(j,q'') : ij \in E(R), q' \in [q_{v,i}^t], q'' \in [q_{v,j}^t]\}.$$

Note that each (i,q') corresponds to the star $E^t_{v,i}(q')$ centred at v. We aim to find a collection of vertex-disjoint cliques of size $k-1$ in R^t_v, which will give us edge-disjoint stars in $E_{G'}(\{v\}, Res_t)$. From the definition, we have

$$|V(R^t_v)| = \sum_{i\in[r]} q^t_{v,i} \overset{\text{(Res2)}}{=} \frac{(1\pm 10\varepsilon)d_{G',Res_t}(v)q}{n/(Tr)} \pm r \overset{(3.6.24)}{=} \frac{(\delta \pm 4d)q|Res_t|}{n/(Tr)} \overset{\text{(Res3)}}{=} (\delta\pm 5d)qr. \quad (3.6.25)$$

Then, for $(i,q')\in V(R^t_v)$, we have

$$\begin{aligned} d_{R^t_v}((i,q')) &\geq q\sum_{j\in N_R(i)} d_{G',V^t_j}(v)|V^t_j|^{-1} - d_R(i) \overset{\text{(Res2)}}{\geq} \frac{Tqr}{(1+7\varepsilon)n}\sum_{j\in N_R(i)} d_{G',V^t_j}(v) - r \\ &\geq \frac{Tqr}{(1+7\varepsilon)n}\Big(\sum_{j\in N_R(i)}|V^t_j| - \sum_{j\in[r]}(|V^t_j| - d_{G',V^t_j}(v))\Big) - r \\ &\overset{(3.6.4),(3.6.24),\text{(Res2)},\text{(Res3)}}{\geq} (2\delta - 2d^{1/2} - 1)qr - r \overset{(3.6.25)}{\geq} (1 - \frac{1}{k-1} + \sigma)|V(R^t_v)|. \end{aligned} \quad (3.6.26)$$

Here, the final inequality follows from (3.3.6). By the Hajnal–Szemerédi theorem, R^t_v contains at least

$$|V(R^t_v)|/(k-1) - 1 \overset{(3.6.25)}{\geq} (\delta - 5d)qr/(k-1) - 1 \overset{\text{(Q2)}}{\geq} (1+d)\kappa \overset{(3.6.23)}{=} \kappa_v$$

vertex-disjoint copies of K_{k-1}. Let $C^t_v(1),\dots,C^t_v(\kappa_v)$ be such vertex-disjoint copies of K_{k-1} in R^t_v. For each $s\in[\kappa_v]$, we let

$$F'_t(v,s) := \bigcup_{(i,q')\in V(C^t_v(s))} E^t_{v,i}(q') \qquad \text{and} \qquad C^t_{v,s} := \{i : (i,q')\in V(C^t_v(s)) \text{ for some } q'\in[q^t_{v,i}]\}.$$

By construction $|C^t_{v,s}| = k-1$ and $R[C^t_{v,s}] \simeq K_{k-1}$. Moreover, the maximum degree of the multi-$(k-1)$-graph $\{C^t_{v,s} : s\in[\kappa_v]\}$ is at most q. Thus we can apply Lemma 3.3.22 with $\{C^t_{v,s} : s\in[\kappa_v]\}$, R, q and k playing the roles of $\mathcal{F}, R, q$ and k. Then we obtain sets $C^{*,t}_{v,s}$ satisfying the following for all $s\in[\kappa_v]$ and $i\in[r]$:

$$C^t_{v,s}\subseteq C^{*,t}_{v,s},\;\; R[C^{*,t}_{v,s}]\simeq K_k,\;\; \text{and}\;\; |\{s\in[\kappa_v] : i\in C^{*,t}_{v,s}\}| \leq (k+1)q. \quad (3.6.27)$$

It is easy to see that for all $s\in[\kappa_v]$ the sets $C^t_{v,s}$, $C^{*,t}_{v,s}$ and the stars $F'_t(v,s)$ satisfy (F′1)–(F′3).

Now, we consider $t \in [T]$ and $v \in V_i \backslash Res_t$ with $i \in [r]$. Let $S_v^t := N_R(i) \backslash \{j : v \in U_i(j)\}$, and for each $j \in S_v^t$, let $E_{v,j}^t$ be a subset of $E_{F_t^*}(\{v\}, V_j^t)$ with $|E_{v,j}^t| = \frac{1}{q}|V_j^t|$. We can choose such a star as there exists $\ell \in \{6, 7\}$ such that

$$d_{F_t^*, V_j^t}(v) = d_{L_\ell, V_j^t}(v) \overset{\text{(L1)}}{=} d \cdot d_{G', V_j^t}(v) \pm n^{2/3} \overset{\text{(Res1),(Res2)}}{=} (1 \pm 10\varepsilon) d \cdot d_{i,j} |V_j^t| > \frac{1}{q}|V_j^t|.$$

Here, the third equality follows since $v \notin U_i(j)$. By (3.6.4), (3.6.5) and the fact that $v \notin U_i'$, we have $|S_v^t| \geq (\delta - 2d^{1/2})r$. Thus

$$\delta(R[S_v^t]) \geq |S_v^t| - (r - \delta(R)) \overset{(3.6.4)}{\geq} (1 - \frac{1}{k-1})|S_v^t|.$$

Again, by the Hajnal–Szemerédi theorem, $R[S_v^t]$ contains (at least) $\kappa_v = \lceil r/(2k) \rceil$ vertex-disjoint copies of K_{k-1}. Denote their vertex sets by $C_{v,1}^t, \dots, C_{v,\kappa_v}^t$. We apply Lemma 3.3.22 with $\{C_{v,s}^t : s \in [\kappa_v]\}$, R, 1 and k playing the roles of $\mathcal{F}, R, q$ and k respectively, to extend each $C_{v,s}^t$ into a $C_{v,s}^{*,t}$ with $R[C_{v,s}^{*,t}] \simeq K_k$ and such that $|\{s \in [\kappa_v] : i \in C_{v,s}^{*,t}\}| \leq k + 1$ for each $i \in [r]$. For each $s \in [\kappa_v]$, let $F_t'(v, s) := \bigcup_{j \in C_{v,s}^t} E_{v,j}^t$. Again, it is easy to see that for all $s \in [\kappa_v]$ the sets $C_{v,s}^t$, $C_{v,s}^{*,t}$ and the stars $F_t'(v, s)$ satisfy (F′1)–(F′3). This proves the claim. □

Altogether we will apply Lemma 3.5.1 κ times in Step 3. In each application, we want the leaves of the stars that we use to be evenly distributed (see condition (A8)$_{3.5.1}$). This will be ensured by Claim 13. More precisely, for each $v \in V(G) \setminus Res_t$, our aim is to choose a permutation $\pi_v^t : [\kappa_v] \to [\kappa_v]$ satisfying the following.

(F′4) For all $t \in [T]$, $i \in [r]$ and $s \in [\kappa/T]$, we have $C(t, s, i) \leq \varepsilon^{4/5} n/r$, where $C(t, s, i) := |\{v \in V_0^{t,s} : i \in C_{v,\pi_v^t(s')}^{*,t}$ for some s' with $(i_t^s(v) - 1)T + 1 \leq s' \leq i_t^s(v)T\}|$,

(F′5) for all $t \in [T]$, $s \in [\kappa/T]$ and $t' \in [T]$, we have that $\bigcup_{v \in V_0^{t,s}} C_{v,\pi_v^t((i_t^s(v)-1)T+t')}^{*,t} \subseteq V(Q_{t,s})$.

Recall from (3.6.20) that $i_t^s(v)$ counts the number of $s' \in [s]$ for which $v \in V_0^{t,s'}$. The number $C(t, s, i)$ is well-defined because $i_t^s(v) \leq \kappa_v/T$ for all $v \in V(G) \setminus Res_t$ by (3.6.21).

Claim 3.6.3. *For each $t \in [T]$ and each $v \in V(G) \setminus Res_t$, there exists a permutation $\pi_v^t : [\kappa_v] \to [\kappa_v]$ satisfying (F′4)–(F′5).*

Proof. We fix $t \in [T]$. We claim that for each $s \in [\kappa/T] \cup \{0\}$ the following hold. For each $v \in V(G) \setminus Res_t$, there exists an injective map $\pi_{v,s}^t : [i_t^s(v)T] \to [\kappa_v]$ satisfying the following.

$(\mathrm{F}'4)_s^t$ For all $i \in [r]$ and $\ell \in [s]$, we have

$$|\{v \in V_0^{t,\ell} : i \in C^{*,t}_{v,\pi^t_{v,s}(s')} \text{ for some } s' \text{ with } (i_t^\ell(v)-1)T+1 \leq s' \leq i_t^\ell(v)T\}| \leq \varepsilon^{4/5}n/r,$$

$(\mathrm{F}'5)_s^t$ for all $\ell \in [s]$ and $t' \in [T]$, we have that $\bigcup_{v \in V_0^{t,\ell}} C^{*,t}_{v,\pi^t_{v,s}((i_t^\ell(v)-1)T+t')} \subseteq V(Q_{t,\ell})$.

Note that both $(\mathrm{F}'4)_0^t$ and $(\mathrm{F}'5)_0^t$ hold by letting $\pi_{v,0}^t : \emptyset \to \emptyset$ be the empty map for all $v \in V(G) \setminus Res_t$. Assume that for some $s \in [\kappa/T-1] \cup \{0\}$ we have already constructed injective maps $\pi^t_{v,s}$ for all $v \in V(G) \setminus Res_t$ which satisfy $(\mathrm{F}'4)_s^t$ and $(\mathrm{F}'5)_s^t$. For each $v \in V_0^{t,s+1}$, we consider the set

$$A_v := \{s' \in [\kappa_v] \setminus \pi^t_{v,s}([i_t^s(v)T]) : C^{*,t}_{v,s'} \subseteq V(Q_{t,s+1})\}.$$

Then we have

$$\begin{aligned} |A_v| &\overset{(\mathrm{F}'2)}{\geq} \kappa_v - i_t^s(v)T - (k+1)q(r - |V(Q_{t,s+1})|) \\ &\overset{(3.6.21),(\mathrm{Q}3)}{\geq} \min\{d \cdot \kappa, r/(2k) - 20T\varepsilon^{1/2}r\} - (k+1)q\varepsilon r \geq r/(4k). \end{aligned} \tag{3.6.28}$$

We choose a subset $I_v \subseteq A_v$ of size T uniformly at random. Then (F′2) implies that for each $i \in V(Q_{t,s+1})$ we have

$$\mathbb{P}[i \in \bigcup_{s' \in I_v} C^{*,t}_{v,s'}] \leq (k+1)qT/|A_v| \leq 10qk^2T/r.$$

Thus

$$\mathbb{E}[|\{v \in V_0^{t,s+1} : i \in \bigcup_{s' \in I_v} C^{*,t}_{v,s'}\}|] \leq 10qk^2T|V_0^{t,s+1}|/r \overset{(3.6.15)}{\leq} \varepsilon^{4/5}n/(2r).$$

A Chernoff bound (Lemma 3.3.1) gives us that for each $i \in V(Q_{t,s+1})$

$$\mathbb{P}\Big[|\{v \in V_0^{t,s+1} : i \in \bigcup_{s' \in I_v} C^{*,t}_{v,s'}\}| \geq \varepsilon^{4/5}n/r\Big] \leq \exp(-\frac{(\varepsilon^{4/5}n/(2r))^2}{2|V_0^{t,s+1}|}) \overset{(3.6.15)}{\leq} e^{-n/r^3}.$$

Since $1 - |V(Q_{t,s+1})|e^{-n/r^3} > 0$, the union bound implies that there exists a choice of I_v for each $v \in V_0^{t,s+1}$ such that for all $i \in V(Q_{t,s+1})$, we have that

$$|\{v \in V_0^{t,s+1} : i \in \bigcup_{s' \in I_v} C^{*,t}_{v,s'}\}| \leq \varepsilon^{4/5} n/r. \tag{3.6.29}$$

If $v \in V(G) \setminus (Res_t \cup V_0^{t,s+1})$ (and thus $i_t^{s+1}(v) = i_t^s(v)$), we let $\pi^t_{v,s+1} := \pi^t_{v,s}$. For each $v \in V_0^{t,s+1}$, we extend $\pi^t_{v,s}$ into $\pi^t_{v,s+1}$ by defining $\pi^t_{v,s+1} : [i_t^{s+1}(v)T] \setminus [i_t^s(v)T] \to I_v$ in an arbitrary injective way. Then, by the choice of I_v, we have that $\pi^t_{v,s+1}$ is an injective map from $[i_t^{s+1}(v)T]$ to $[\kappa_v]$ satisfying $(\mathrm{F}'5)^t_{s+1}$. Moreover, (3.6.29) implies that for any $i \in V(Q_{t,s+1})$, we have

$$\begin{aligned}
&|\{v \in V_0^{t,s+1} : i \in C^{*,t}_{v,\pi^t_{v,s+1}(s')} \text{ for some } s' \text{ with } (i_t^{s+1}(v) - 1)T + 1 \leq s' \leq i_t^{s+1}(v)T\}| \\
&= |\{v \in V_0^{t,s+1} : i \in \bigcup_{s' \in I_v} C^{*,t}_{v,s'}\}| \overset{(3.6.29)}{\leq} \varepsilon^{4/5} n/r.
\end{aligned}$$

This with $(\mathrm{F}'4)^t_s$ implies $(\mathrm{F}'4)^t_{s+1}$. By repeating this, we obtain injective maps $\pi^t_{v,\kappa/T}$ satisfying both $(\mathrm{F}'4)^t_{\kappa/T}$ and $(\mathrm{F}'5)^t_{\kappa/T}$. For each $v \in V(G) \setminus Res_t$, we extend $\pi^t_{v,\kappa/T}$ into a permutation $\pi^t_v : [\kappa_v] \to [\kappa_v]$ by assigning arbitrary values for the remaining values in the domain. It is easy to see that $(\mathrm{F}'4)^t_{\kappa/T}$ implies (F′4) and $(\mathrm{F}'5)^t_{\kappa/T}$ implies (F′5). We can find such permutations for all $t \in [T]$. Thus such collection satisfies both (F′4) and (F′5). □

For each $t \in [T]$, let

$$G_t := G_t^* \cup \bigcup_{s \in [\kappa/T]} G_{t,s} \quad \text{and} \quad F'_t := \bigcup_{v \in V(G) \setminus Res_t} \bigcup_{s \in [\kappa_v]} F'_t(v, s).$$

Then $G_1, \ldots, G_T, F_1, \ldots, F_T, F'_1, \ldots, F'_T$ form edge-disjoint subgraphs of G. (Recall that G_t^* was defined in (3.6.8), $G_{t,s}$ in (3.6.10) and $F'_v(t, s)$ in Claim 3.6.2.)

Step 2. Partitioning $\mathcal{H}$. Now we will partition $\mathcal{H}$. Recall that the graphs in $\mathcal{H}$ are η^2-separable. By packing several graphs from $\mathcal{H}$ with less than $n/4$ edges suitably into a single graph in a way that no edges from distinct graphs intersect each other, we can assume that all but at most one graph in $\mathcal{H}$ have at least $n/4$ edges, and that all graphs in $\mathcal{H}$ are (k,η)-chromatic, η-separable and have maximum degree at most Δ. By adding at most $n/4$ edges to at most one graph if necessary, we can then assume that all graphs in $\mathcal{H}$ have at least $n/4$ edges. Moreover, if $e(\mathcal{H})$ is too small, we can add some copies of n-vertex paths to $\mathcal{H}$ to assume that

$$\varepsilon n^2 \leq e(\mathcal{H}) \overset{(iii)}{\leq} (1-\nu)e(G) + n/4.$$

We partition $\mathcal{H}$ into κ collections $\mathcal{H}_{1,1},\dots,\mathcal{H}_{T,\kappa/T}$ such that for all $t \in [T]$ and $s \in [\kappa/T]$, we have

$$n^{7/4} \overset{\text{(Q2)}}{\leq} \frac{\varepsilon n^2}{\kappa} - \Delta n \leq e(\mathcal{H}_{t,s}) < \frac{1}{\kappa}(1-\nu)e(G) + 2\Delta n \overset{(i),\text{(Q2)}}{\leq} \frac{(1-2\nu/3)(k-1)n^2}{2qr}. \quad (3.6.30)$$

Indeed, this is possible since $e(H) \leq \Delta n$ for all $H \in \mathcal{H}$. Now, we are ready to construct the desired packing.

Step 3. Construction of packings into the host graphs. As $G_1,\dots,G_T,F_1,\dots,F_T,F'_1,\dots,F'_T$ are edge-disjoint subgraphs of G, and $\mathcal{H}_{1,1},\dots,\mathcal{H}_{T,\kappa/T}$ is a partition of $\mathcal{H}$, it suffices to show that for each $t \in [T]$, we can pack $\mathcal{H}_t := \bigcup_{s=1}^{\kappa/T} \mathcal{H}_{t,s}$ into $G_t \cup \bigcup_{s\in[\kappa/T]} F_{t,s} \cup F'_t$. (Recall from (3.6.9) that $F_{t,1},\dots,F_{t,\kappa/T}$ are edge-disjoint subgraphs of F_t.) We fix $t \in [T]$ and will apply Lemma 3.5.1 κ/T times to show that such a packing exists.

Assume that for some s with $0 \leq s \leq \kappa/T - 1$, we have already defined a function ϕ_s packing $\bigcup_{s'=1}^{s} \mathcal{H}_{t,s'}$ into $G_t \cup F_t \cup F'_t$ and satisfying the following, where $\Phi^s := \bigcup_{s'=1}^{s} \phi_s(\mathcal{H}_{t,s'})$ and $j_t^s(u)$ is defined in (3.6.20) and G_t^* is defined in (3.6.8).

(G1)$_s$ For each $u \in Res_t$, we have $d_{\Phi^s \cap G_t^*}(u) \leq \frac{4k\Delta j_t^s(u) n}{qr} + \frac{\varepsilon^{1/9} sn}{r}$,

(G2)$_s$ for each $i \in [r]$, we have $e_{\Phi^s \cap G_t^*}(V_i \backslash V_i^t, Res_t) \leq \frac{\varepsilon^{1/3} sn^2}{r^2}$,

(G3)$_s$ for $s' \in [\kappa/T] \setminus [s]$, we have $E(\Phi^s) \cap (E(G_{t,s'}) \cup E(F_{t,s'})) = \emptyset$,

(G4)$_s$ for $v \in V(G) \setminus Res_t$, $s'' \in [\kappa_v]$ with $s'' > i_t^s(v) \cdot T$, we have $E(\Phi^s) \cap F'_t(v, \pi_v^t(s'')) = \emptyset$.

Note that (G1)$_0$–(G4)$_0$ trivially hold with an empty packing $\phi_0 : \emptyset \to \emptyset$. For each $t' \in [T]$ and $v \in V(G) \setminus Res_t$, let $\ell(v,t') := \pi^t_v((i^{s+1}_t(v)-1)T+t')$. (Note that $\ell(v,t')$ is well-defined since $(i^{s+1}_t(v)-1)T+t' \leq \kappa_v$ by (3.6.21).) Let

$$V := \bigcup_{i \in V(Q_{t,s+1})} V^{t,s+1}_i, \quad U := \bigcup_{i \in V(Q_{t,s+1})} U^{t,s+1}_i, \tag{3.6.31}$$

$$\hat{G} := (G_{t,s+1}[V] \cup G^*_t[V \cup Res_t]) \setminus E(\Phi^s), \text{ and } \hat{F}' := \bigcup_{v \in V^{t,s+1}_0} \bigcup_{t' \in [T]} F'_t(v, \ell(v,t'))[\{v\}, U]. \tag{3.6.32}$$

Note that (G3)$_s$ implies that $E(F_{t,s+1}) \cap E(\Phi^s) = \emptyset$. Let $\hat{R}$ be the graph on vertex set $V(Q_{t,s+1})$ with

$$E(\hat{R}) := \{ij \in E(R[V(Q_{t,s+1})]) : |E_{G^*_t}(V_i, V_j) \cap E(\Phi^s)| < \varepsilon^{1/10} n^2/r^2\}.$$

We wish to apply Lemma 3.5.1 with the following objects and parameters.

object/parameter	$\hat{G}$	$F_{t,s+1}[V,U]$	$\hat{F}'$	$V^{t,s+1}_0$	$U^{t,s+1}_0$	$V^{t,s+1}_i$	$U^{t,s+1}_i$	$\hat{R}$
playing the role of	G	F	F'	V_0	U_0	V_i	U_i	R
object/parameter	$1/q$	$\mathcal{H}_{t,s+1}$	d	$C^{*,t}_{v,\ell(v,t')}$	$C^t_{v,\ell(v,t')}$	$F'_t(v,\ell(v,t'))[\{v\},U]$	k	Δ
playing the role of	α	$\mathcal{H}$	d	$C^*_{v,t}$	$C_{v,t}$	$F'_{v,t}$	k	Δ
object/parameter	$Q_{t,s+1}$	η	25ε	$\sigma/2$	T	$\nu/2$	m	
playing the role of	Q	η	ε	σ	T	ν	n'	

Thus $Res_t \setminus U^{t,s+1}_0$ plays the role of $U = \bigcup_{i=1}^r U_i$ in Lemma 3.5.1, and $t' \in [T]$ stands for $t \in [T]$. By (3.6.1), (3.6.14), (3.6.15), (3.6.16), (Q3) and (F′5) we have appropriate objects and parameters as well as the hierarchy of constants required in Lemma 3.5.1. Now we show that (A1)$_{3.5.1}$–(A9)$_{3.5.1}$ hold. (A1)$_{3.5.1}$ is obvious from Theorem 3.1.2 (ii) and our assumption in Step 2. (A2)$_{3.5.1}$ holds by (3.6.30). (A3)$_{3.5.1}$ follows from Claim 3.6.1 and (G3)$_s$. Consider $ij \in E(\hat{R})$, then $\hat{G}[U^{t,s+1}_i, U^{t,s+1}_j] = G^*_t[U^{t,s+1}_i, U^{t,s+1}_j] \setminus E(\Phi^s)$. Since $U^{t,s+1}_i \subseteq V_i$ and

$U_j^{t,s+1} \subseteq V_j$, the properties (3.6.16), (3.6.17) and the definition of $\hat{R}$ imply that

$$e_{G_t^*}(U_i^{t,s+1}, U_j^{t,s+1}) - e_{\Phi^s \cap G_t^*}(V_i, V_j) \geq (1 - \varepsilon^{1/15}) e_{G_t^*}(U_i^{t,s+1}, U_j^{t,s+1}).$$

Thus, Proposition 3.3.3 with (3.6.17) implies that $\hat{G}[U_i^{t,s+1}, U_j^{t,s+1}]$ is $(\varepsilon^{1/50}, (d^2))^+$-regular. The calculation for $\hat{G}[V_i^{t,s+1}, U_j^{t,s+1}]$ is similar. Thus (A4)$_{3.5.1}$ holds with the above objects and parameters. By (G1)$_s$, for each $i \in [r]$ we have

$$e_{\phi^s \cap G_t^*}(V_i^t, \bigcup_{j \in [r] \setminus \{i\}} V_j) \leq \sum_{v \in V_i^t} \Big(\frac{4k\Delta j_t^s(v) n}{qr} + \frac{\varepsilon^{1/9} sn}{r} \Big) \overset{\text{(Q2)},(3.6.22),\text{(Res2)}}{\leq} \frac{\varepsilon^{1/9} n^2}{r}. \quad (3.6.33)$$

Thus, for $i \in V(Q_{t,s+1}) = V(\hat{R})$, we have

$$\begin{aligned} d_R(i) - d_{\hat{R}}(i) &\leq \frac{e_{\Phi^s \cap G_t^*}(V_i \setminus V_i^t, Res_t) + e_{\Phi^s \cap G_t^*}(V_i^t, \bigcup_{j \in [r] \setminus \{i\}} V_j)}{\varepsilon^{1/10} n^2 / r^2} + |V(R) \setminus V(\hat{R})| \\ &\overset{\text{(G2)}_s,\text{(Q3)},(3.6.33)}{\leq} \frac{\varepsilon^{1/3} sn^2/r^2 + \varepsilon^{1/9} n^2/r}{\varepsilon^{1/10} n^2/r^2} + \varepsilon r \overset{\text{(Q2)}}{\leq} \varepsilon^{1/100} r. \end{aligned}$$

This with (3.6.4) and (3.3.6) implies that (A5)$_{3.5.1}$ holds for $\hat{R}$. For all $ij \in E(Q_{t,s+1})$ and $u \in U_i^{t,s+1}$, by (3.6.18), we have

$$d_{F_{t,s+1}, V_j^{t,s+1}}(u) \geq 2d^2 |V_j \setminus Res_t|/3 \overset{\text{(Res2)},(3.6.14)}{\geq} d^3 m.$$

Thus (A6)$_{3.5.1}$ holds. By (F′1), (F′4) and the fact that $i_t^{s+1}(v) = i_t^s(v) + 1$ for all $v \in V_0^{t,s}$, (A8)$_{3.5.1}$ holds (for $C^{*,t}_{v,\ell(v,t')}$, $C^t_{v,\ell(v,t')}$ and all $v \in V_0^{t,s}$). If $v \in V_0^{t,s}, t' \in [T]$ and $i \in C^t_{v,\ell(v,t')} \subseteq C^{*,t}_{v,\ell(v,t')}$ then (F′5) implies that $i \in V(Q_{t,s+1})$. Moreover, by (3.6.16) we have $|U_i^{t,s+1}| \geq |V_i^t| - 5k\varepsilon^2 n/r$. Together with (F′3) this implies that $d_{F'_{t'}(v,\ell(v,t')), U_i^{t,s+1}}(v) \geq (1-\varepsilon)|U_i^{t,s+1}|/q$. Thus (A7)$_{3.5.1}$ holds. To check (A9)$_{3.5.1}$, note that for each $u \in U_0^{t,s+1}$, we have

$$d_{G_t^* \cap \Phi^s}(u) \overset{\text{(G1)}_s}{\leq} 4k\Delta j_t^s(u) n/(qr) + \varepsilon^{1/9} sn/r \overset{\text{(Q2)},(3.6.22)}{\leq} \varepsilon^{1/10} n.$$

Thus,

$$\begin{aligned} |\{i \in V(Q_{t,s+1}) : d_{\hat{G}, V_i^{t,s+1}}(u) \geq d^2 m/3\}| &\geq |\{i \in V(Q_{t,s+1}) : d_{G_t^*, V_i^{t,s+1}}(u) \geq d^2 m/2\}| - \frac{d_{G_t^* \cap \Phi^s}(u)}{d^2 m/6} \\ &\overset{(3.6.19)}{\geq} (1 - 1/k + \sigma/2) r - \frac{\varepsilon^{1/10} n}{d^2 m/6} \overset{(3.6.14)}{\geq} (1 - 1/k + \sigma/3) r. \end{aligned}$$

This implies that

$$|\{i \in V(Q_{t,s+1}) : d_{\hat{G},V_j^{t,s+1}}(u) \geq d^3 m \text{ for all } j \in N_{Q_{t,s+1}}(i)\}| \geq \sigma^2 r.$$

This shows that $(\mathrm{A9})_{3.5.1}$ holds. Hence, by Lemma 3.5.1, we obtain a function ψ_{s+1} packing $\mathcal{H}_{t,s+1}$ into $\hat{G} \cup F_{t,s+1} \cup \hat{F}'$ and satisfying the following.

(B1) $\Delta(\psi_{s+1}(\mathcal{H}_{t,s+1})) \leq 4k\Delta n/(qr)$,

(B2) for each $u \in Res_t \setminus U_0^{t,s+1}$, we have $d_{\psi_{s+1}(\mathcal{H}_{t,s+1})\cap\hat{G}}(u) \leq 10\Delta\varepsilon^{1/8}n/r$,

(B3) for each $i \in V(Q_{t,s})$, we have $e_{\psi_{s+1}(\mathcal{H}_{t,s+1})\cap\hat{G}}(V_i^{t,s+1}, Res_t) < 10\varepsilon^{1/2}n^2/r^2$.

Moreover, $(\mathrm{G3})_s$ with $(\mathrm{G4})_s$ implies that $\psi_{s+1}(\mathcal{H}_{t,s+1})$ is edge-disjoint from Φ^s, thus the map $\phi_{s+1} := \phi_s \cup \psi_{s+1}$ packs $\bigcup_{s'=1}^{s+1} \mathcal{H}_{t,s'}$ into $G_t \cup \bigcup_{s'=1}^{\kappa/T} F_{t,s'} \cup F'_t$. Now it remains to show that ϕ_{s+1} satisfies $(\mathrm{G1})_{s+1}$–$(\mathrm{G4})_{s+1}$.

Consider any vertex $u \in Res_t$. If $u \in U_0^{t,s+1}$, then we know that $j_t^{s+1}(u) = j_t^s(u) + 1$. Thus $(\mathrm{G1})_s$ together with (B1) implies $(\mathrm{G1})_{s+1}$ for the vertex u. If $u \in Res_t \setminus U_0^{t,s+1}$, then we have $j_t^{s+1}(u) = j_t^s(u)$, thus $(\mathrm{G1})_s$ together with (B2) implies $(\mathrm{G1})_{s+1}$.

For each $i \in [r]$, (3.6.31) implies that the vertices in $V_i \setminus (V_i^{t,s+1} \cup V_i^t) \subseteq V_0^{t,s+1}$ are not incident to any edges in $\Phi^{s+1} \cap G_t^*$. Thus it is easy to see that $(\mathrm{G2})_s$ together with (B3) implies $(\mathrm{G2})_{s+1}$. As ψ_{s+1} packs $\mathcal{H}_{t,s+1}$ into $\hat{G} \cup F_{t,s+1} \cup \hat{F}'$, (3.6.32) together with $(\mathrm{G3})_s$ implies $(\mathrm{G3})_{s+1}$. Moreover, we have

$$i_t^{s+1}(v) = \begin{cases} i_t^s(v) + 1 & \text{if } v \in V_0^{t,s+1}, \\ i_t^s(v) & \text{otherwise.} \end{cases}$$

Thus, (3.6.32) together with $(\mathrm{G4})_s$ and the definition of $\ell(v,t')$ implies $(\mathrm{G4})_{s+1}$.

By repeating this for each $s \in [\kappa/T]$ in order, we obtain a function $\phi_{\kappa/T}$ which packs $\mathcal{H}_t$ into $G_t \cup F_t \cup F'_t$. By taking the union of such functions over all $t \in [T]$, we obtain a desired function packing $\mathcal{H}$ into $\bigcup_{t\in[T]} G_t \cup F_t \cup F'_t \subseteq G$. This completes the proof. □

The proof of Theorem 3.1.5, follows almost exactly the same lines as that of Theorem 3.1.2, with one very minor difference. Indeed, the only place where we need the condition that G is almost regular is when we apply Lemma 3.3.13 in Step 1 to obtain (Q1)–(Q5). Thus to prove Theorem 3.1.5, we only need to replace the application of Lemma 3.3.13 with an application of the following result. (Note that (B1) below implies both (Q3) and (Q4).)

Lemma 3.6.1. *Suppose $n, q, T \in \mathbb{N}$ with $0 < 1/n \ll \varepsilon, 1/T, 1/q, \nu \leq 1/2$ and $0 < 1/n \ll \nu < \sigma/2 < 1$ and $\delta = 1/2 + \sigma$ and q divides T. Let G be an n-vertex multi-graph with edge-multiplicity at most q, such that for all $v \in V(G)$ we have $d_G(v) \geq q\delta n$.*

Then there exists a subset $V' \subseteq V(G)$ with $|V'| \leq 1$ and $|V(G)\backslash V'|$ being even, and there exist pairwise edge-disjoint matchings $F_{1,1}, \ldots, F_{1,\kappa}, F_{2,1}, \ldots, F_{T,\kappa}$ of G with $\kappa = \frac{(\delta+\sqrt{2\delta-1}-\nu)qn}{2T} \pm 1$ satisfying the following.

(B1) *For each $(t', i) \in [T] \times [\kappa]$, we have that $V(F_{t',i}) = V(G)\backslash V'$,*

(B2) *for all $t' \in [T]$ and $u, v \in V(G)$, we have $|\{i \in [\kappa] : u \in N_{F_{t',i}}(v)\}| \leq 1$.*

The proof of the above lemma is very similar (but simpler) than that of Lemma 3.3.13. We proceed as in the proof of Lemma 3.3.13 to obtain simple graphs G^c with $\delta(G^c) > \delta n - \nu^2 n$. We let $V' \subseteq V(G)$ be such that $|V'| \leq 1$ and $|V(G)\backslash V'|$ is even. The difference is that we now apply the following result of [30] to each $G^c_* := G^c[V(G)\backslash V']$ to obtain the desired matchings M^c_i: for every $\alpha > 0$, any sufficiently large n-vertex graph with minimum degree $\delta \geq (1/2 + \alpha)n$ contains at least $(\delta - \alpha n + \sqrt{n(2\delta - n)})/4$ edge-disjoint Hamilton cycles.

Acknowledgement

We are grateful to the referee for helpful comments on an earlier version.

References

[1] S. Abbasi. *The solution of the El-Zahar problem.* Ph.D. Thesis, Rutgers University, 1998.

[2] M. Ajtai, J. Komlós, and E. Szemerédi. “First occurrence of Hamilton cycles in random graphs”. In: *Cycles in graphs (Burnaby, B.C., 1982).* Vol. 115. North-Holland Math. Stud. North-Holland, Amsterdam, 1985, pp. 173–178. DOI: 10.1016/S0304-0208(08)73007-X.

[3] P. Allen, J. Böttcher, J. Hladkỳ, and D. Piguet. “Packing degenerate graphs”. In: *Adv. Math.* 354 (2019). DOI: 10.1016/j.aim.2019.106739.

[4] N. Alon, P. Seymour, and R. Thomas. “A separator theorem for nonplanar graphs”. In: *J. Amer. Math. Soc.* 3 (1990), pp. 801–808. DOI: 10.1090/S0894-0347-1990-1065053-0.

[5] N. Alon and J. H. Spencer. *The probabilistic method.* Fourth. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, 2016.

[6] Y. Alon and M. Krivelevich. “Hitting time of edge disjoint Hamilton cycles in random subgraph processes on dense base graphs”. In: *arXiv e-prints* (2019). eprint: 1912.01251.

[7] J. Balogh, A. Treglown, and A. Z. Wagner. “Tilings in randomly perturbed dense graphs”. In: *Combin. Probab. Comput.* 28 (2019), pp. 159–176. DOI: 10.1017/S0963548318000366.

[8] B. Barber, D. Kühn, A. Lo, and D. Osthus. “Edge-decompositions of graphs with high minimum degree”. In: *Adv. Math.* 288 (2014), pp. 337–385. DOI: 10.1016/j.aim.2015.09.032.

[9] S. Bhatt and J.-Y. Cai. “Take a walk, grow a tree”. In: *29th Annual Symposium on Foundations of Computer Science (STOC)*. 1988, pp. 469–478.

[10] S. N. Bhatt, F. R. K. Chung, F. T. Leighton, and A. L. Rosenberg. “Efficient embeddings of trees in hypercubes”. In: *SIAM J. Comput.* 21 (1992), pp. 151–162. DOI: 10.1137/0221012.

[11] T. Bohman, A. Frieze, and R. Martin. “How many random edges make a dense graph Hamiltonian?” In: *Random Structures Algorithms* 22 (2003), pp. 33–42. DOI: 10.1002/rsa.10070.

[12] B. Bollobás, Y. Kohayakawa, and T. Łuczak. “The evolution of random subgraphs of the cube”. In: *Random Structures Algorithms* 3 (1992), pp. 55–90. DOI: 10.1002/rsa.3240030106.

[13] B. Bollobás. “The evolution of the cube”. In: *Combinatorial mathematics (Marseille-Luminy, 1981)*. Vol. 75. North-Holland Math. Stud. North-Holland, Amsterdam, 1983, pp. 91–97.

[14] B. Bollobás. “The evolution of sparse graphs”. In: *Graph theory and combinatorics (Cambridge, 1983)*. Academic Press, London, 1984, pp. 35–57.

[15] B. Bollobás. “Complete matchings in random subgraphs of the cube”. In: *Random Structures Algorithms* 1 (1990), pp. 95–104. DOI: 10.1002/rsa.3240010107.

[16] B. Bollobás. Personal communication. 2020.

[17] B. Bollobás and A. M. Frieze. “On matchings and Hamiltonian cycles in random graphs”. In: *Random graphs ’83 (Poznań, 1983)*. Vol. 118. North-Holland Math. Stud. North-Holland, Amsterdam, 1985, pp. 23–46.

[18] B. Bollobás and A. Thomason. “Random graphs of small order”. In: *Random graphs ’83 (Poznań, 1983)*. Vol. 118. North-Holland Math. Stud. North-Holland, Amsterdam, 1985, pp. 47–97.

[19] C. Borgs, J. T. Chayes, R. van der Hofstad, G. Slade, and J. Spencer. "Random subgraphs of finite graphs. III. The phase transition for the n-cube". In: *Combinatorica* 26 (2006), pp. 395–410. DOI: 10.1007/s00493-006-0022-1.

[20] J. Böttcher, J. Han, Y. Kohayakawa, R. Montgomery, O. Parczyk, and Y. Person. "Universality for bounded degree spanning trees in randomly perturbed graphs". In: *Random Structures Algorithms* 55 (2019), pp. 854–864. DOI: 10.1002/rsa.20850.

[21] J. Böttcher, J. Hladkỳ, D. Piguet, and A. Taraz. "An approximate version of the Tree Packing Conjecture". In: *Israel J. Math.* 211 (2014). DOI: 10.1007/s11856-015-1277-2.

[22] J. Böttcher, R. Montgomery, O. Parczyk, and Y. Person. "Embedding spanning bounded degree graphs in randomly perturbed graphs". In: *Mathematika* 66 (2020), pp. 422–447. DOI: 10.1112/mtk.12005.

[23] J. Böttcher, K. Pruessmann, A. Taraz, and A. Würfl. "Bandwidth, expansion, treewidth, separators and universality for bounded-degree graphs". In: *Eur. J. Combin.* 31 (2010), pp. 1217–1227. DOI: 10.1016/j.ejc.2009.10.010.

[24] J. Böttcher, M. Schacht, and A. Taraz. "Proof of the bandwidth conjecture of Bollobás and Komlós". In: *Math Ann* 343 (2009), pp. 175–205. DOI: 10.1007/s00208-008-0268-6.

[25] D. Bryant, D. Horsley, and W. Pettersson. "Cycle decompositions V: Complete graphs into cycles of arbitrary lengths". In: *Proc. London Math. Soc.* 108 (2012), pp. 1153–1192. DOI: 10.1112/plms/pdt051.

[26] J. D. Burtin. "The probability of connectedness of a random subgraph of an n-dimensional cube". In: *Problemy Peredači Informacii* 13 (1977), pp. 90–95.

[27] R. Caha and V. Koubek. "Spanning multi-paths in hypercubes". In: *Discrete Math.* 307 (2007), pp. 2053–2066. DOI: 10.1016/j.disc.2005.12.050.

[28] M. Y. Chan and S.-J. Lee. "On the existence of Hamiltonian circuits in faulty hypercubes". In: *SIAM J. Discrete Math.* 4 (1991), pp. 511–527. DOI: 10.1137/0404045.

[29] X.-B. Chen. “Paired many-to-many disjoint path covers of the hypercubes”. In: *Inform. Sci.* 236 (2013), pp. 218–223. DOI: 10.1016/j.ins.2013.02.028.

[30] D. Christofides, D. Kühn, and D. Osthus. “Edge-disjoint Hamilton cycles in graphs”. In: *J. Combin. Theory Ser. B* 102 (2009), pp. 1035–1060. DOI: 10.1016/j.jctb.2011.10.005.

[31] P. Condon, A. Espuny Díaz, A. Girão, D. Kühn, and D. Osthus. "Hamiltonicity of random subgraphs of the hypercube". Submitted for publication.

[32] P. Condon, A. Espuny Díaz, A. Girão, D. Kühn, and D. Osthus. “Dirac's theorem for random regular graphs”. In: *arXiv e-prints* (2018). arXiv: 1903.05052.

[33] P. Condon, A. Espuny Díaz, J. Kim, D. Kühn, and D. Osthus. “Resilient Degree Sequences with respect to Hamilton Cycles and Matchings in Random Graphs”. In: *Electron. J. Combin.* 26 (2019), P4.54. DOI: 10.37236/8279.

[34] B. Csaba, I. Levitt, J. Nagy-György, and E. Szemerédi. *Fete of Combinatorics, Bolyai Society Mathematical Studies.* Vol. 20. Springer, Berlin, Heidelberg, 2010, pp. 95–137.

[35] B. Csaba, D. Kühn, A. Lo, D. Osthus, and A. Treglown. “Proof of the 1-factorization and Hamilton Decomposition Conjectures”. In: *Mem. Amer. Math. Soc.* 244 (2016), monograph 1154, 170 pages. DOI: 10.1090/memo/1154.

[36] B. Csaba, A. Shokoufandeh, and E. Szemerédi. “Proof of a Conjecture of Bollobás and Eldridge for Graphs of Maximum Degree Three”. In: *Combinatorica* 23 (2003), pp. 35–72. DOI: 10.1007/s00493-003-0013-4.

[37] F. Dross. “Fractional Triangle Decompositions in Graphs with Large Minimum Degree”. In: *SIAM J. Discrete Math.* 30 (2015), pp. 36–42. DOI: 10.1137/15M1014310.

[38] P. Dukes and A. Ling. “Asymptotic Existence of Resolvable Graph Designs”. In: *Canadian Mathematical Bulletin* 50 (2007), pp. 504–518. DOI: 10.4153/CMB-2007-050-x.

[39] T. Dvořák and P. Gregor. "Partitions of faulty hypercubes into paths with prescribed endvertices". In: *SIAM J. Discrete Math.* 22 (2008), pp. 1448–1461. DOI: 10.1137/060678476.

[40] T. Dvořák, P. Gregor, and V. Koubek. "Generalized Gray codes with prescribed ends". In: *Theoret. Comput. Sci.* 668 (2017), pp. 70–94. DOI: 10.1016/j.tcs.2017.01.010.

[41] M. E. Dyer, A. M. Frieze, and L. R. Foulds. "On the strength of connectivity of random subgraphs of the n-cube". In: *Random graphs '85 (Poznań, 1985)*. Vol. 144. North-Holland Math. Stud. North-Holland, Amsterdam, 1987, pp. 17–40.

[42] P. Erdős and J. Spencer. "Evolution of the n-cube". In: *Comput. Math. Appl.* 5 (1979), pp. 33–39. DOI: 10.1016/0898-1221(81)90137-1.

[43] A. Ferber, C. Lee, and F. Mousset. "Packing spanning graphs from separable families". In: *Israel J. Math.* 219 (2015), pp. 959–982. DOI: 10.1007/s11856-017-1504-0.

[44] A. Ferber and W. Samotij. "Packing trees of unbounded degrees in random graphs". In: *J. London Math. Soc.* (2016), pp. 653–677. DOI: 10.1112/jlms.12179.

[45] J. A. Fill and R. Pemantle. "Percolation, first-passage percolation and covering times for Richardson's model on the n-cube". In: *Ann. Appl. Probab.* 3 (1993), pp. 593–629.

[46] A. Frieze. "Random structures and algorithms". In: *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. 1.* Kyung Moon Sa, Seoul, 2014, pp. 311–340.

[47] A. Frieze. "Hamilton cycles in random graphs: a bibliography". In: *arXiv e-prints* (2019). arXiv: 1901.07139.

[48] S. Glock, F. Joos, J. Kim, D. Kühn, and D. Osthus. *Resolution of the Oberwolfach problem*. 2018. arXiv: 1806.04644.

[49] S. Glock, D. Kühn, A. Lo, R. Montgomery, and D. Osthus. “On the decomposition threshold of a given graph”. In: *J. Combin. Theory Ser. B* (2016). DOI: 10.1016/j.jctb.2019.02.010.

[50] S. Glock, D. Kühn, A. Lo, and D. Osthus. “Hypergraph F-designs for arbitrary F”. In: *arXiv e-prints* (2019). arXiv: 1706.01800.

[51] S. Glock, D. Kühn, A. Lo, and D. Osthus. “The existence of designs via iterative absorption”. In: *arXiv e-prints* (2019). arXiv: 1611.06827.

[52] P. Gregor and T. Dvořák. “Path partitions of hypercubes”. In: *Inform. Process. Lett.* 108 (2008), pp. 402–406. DOI: 10.1016/j.ipl.2008.07.015.

[53] M. Hahn-Klimroth, G. S. Maesaka, Y. Mogge, S. Mohr, and O. Parczyk. “Random perturbation of sparse graphs”. In: *arXiv e-prints* (2020). arXiv: 2004.04672.

[54] A. Hajnal and E. Szemerédi. *Proof of a conjecture of P. Erdős.* In Combinatorial theory and its applications, II (Proc. Colloq., Balatonfüred, 1969), 1970.

[55] P. Haxell and V. Rödl. “Integer and Fractional Packings in Dense Graphs”. In: *Combinatorica* 21 (2001), pp. 13–38. DOI: 10.1007/s004930170003.

[56] R. van der Hofstad and A. Nachmias. “Hypercube percolation”. In: *J. Eur. Math. Soc. (JEMS)* 19 (2017), pp. 725–814. ISSN: 1435-9855. DOI: 10.4171/JEMS/679.

[57] A. Jamshed. *Embedding spanning subgraphs into large dense graphs.* Ph.D. Thesis, Rutgers University, 2010.

[58] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000. DOI: 10.1002/9781118032718.

[59] T. Johansson. “On Hamilton cycles in Erdős-Rényi subgraphs of large graphs”. In: *arXiv e-prints* (2018). eprint: 1811.03501.

[60] F. Joos, J. Kim, D. Kühn, and D. Osthus. "Optimal packings of bounded degree trees". In: *J. Eur. Math. Soc. (JEMS)* (2016), pp. 3573–3647. DOI: 10.4171/JEMS/909.

[61] H. Kaul, A. Kostochka, and G. Yu. "On a graph packing conjecture by Bollobás, Eldridge and Catlin". In: *Combinatorica* 28 (2008), pp. 469–485. DOI: 10.1007/s00493-008-2278-0.

[62] P. Keevash. "The existence of designs". In: *arXiv e-prints* (Jan. 2014). arXiv: 1401.3665.

[63] J. Kim, D. Kühn, D. Osthus, and M. Tyomkyn. "A blow-up lemma for approximate decompositions". In: *Trans. Amer. Math. Soc.* 371 (2016), pp. 4655–4742. DOI: 10.1090/tran/7411.

[64] F. Knox, D. Kühn, and D. Osthus. "Edge-disjoint Hamilton cycles in random graphs". In: *Random Structures Algorithms* 46 (2015), pp. 397–445. DOI: 10.1002/rsa.20510.

[65] D. E. Knuth. *The Art of Computer Programming, Volume 4, Fascicle 2: Generating All Tuples and Permutations (Art of Computer Programming)*. Addison-Wesley Professional, 2005.

[66] Y. Kohayakawa, B. Kreuter, and D. Osthus. "The length of random subsets of Boolean lattices". In: *Random Structures Algorithms* 16 (2000), pp. 177–194. DOI: 10.1002/(SICI)1098-2418(200003)16:2<177::AID-RSA4>3.0.CO;2-9.

[67] J. Komlós, G. Sárközy, and E. Szemerédi. "Proof of a packing conjecture of Bollobás". In: *Combin. Probab. Comput.* 4 (1995), pp. 241–255.

[68] J. Komlós, G. Sárközy, and E. Szemerédi. "Blow-up lemma". In: *Combinatorica* 17 (1997), pp. 109–123.

[69] J. Komlós, G. Sárközy, and E. Szemerédi. "On the Pósa-Seymour conjecture". In: *J. Graph Theory* 29 (1998), pp. 167–176.

[70] J. Komlós, G. Sárközy, and E. Szemerédi. "Proof of the Seymour conjecture for large graphs". In: *Ann. Combin.* 2 (1998), pp. 43–60.

[71] J. Komlós, G. Sárközy, and E. Szemerédi. "Spanning trees in dense graphs". In: *Combin. Probab. Comput.* 10 (2001), pp. 397–416.

[72] J. Komlós and E. Szemerédi. "Limit distribution for the existence of Hamiltonian cycles in a random graph". In: *Discrete Math.* 43 (1983), pp. 55–63. DOI: 10.1016/0012-365X(83)90021-3.

[73] A. D. Koršunov. "Solution of a problem of P. Erdős and A. Rényi on Hamiltonian cycles in undirected graphs". In: *Dokl. Akad. Nauk SSSR* 228 (1976), pp. 529–532.

[74] M. Krivelevich, M. Kwan, and B. Sudakov. "Bounded-degree spanning trees in randomly perturbed graphs". In: *SIAM J. Discrete Math.* 31 (2017), pp. 155–171. DOI: 10.1137/15M1032910.

[75] M. Krivelevich, C. Lee, and B. Sudakov. "Robust Hamiltonicity of Dirac graphs". In: *Trans. Amer. Math. Soc.* 366 (2014), pp. 3095–3130. DOI: 10.1090/S0002-9947-2014-05963-1.

[76] M. Krivelevich and W. Samotij. "Optimal packings of Hamilton cycles in sparse random graphs". In: *SIAM J. Discrete Math.* 26 (2012), pp. 964–982. DOI: 10.1137/110849171.

[77] D. Kühn and D. Osthus. "Embedding large subgraphs into dense graphs". In: *London Math. Soc. Lecture Notes* 365 (2009), pp. 137–167.

[78] D. Kühn and D. Osthus. "The minimum degree threshold for perfect graph packings". In: *Combinatorica* 29 (2009), pp. 65–107. DOI: 10.1007/s00493-009-2254-3.

[79] D. Kühn and D. Osthus. "Hamilton decompositions of regular expanders: A proof of Kelly's conjecture for large tournaments". In: *Adv. Math.* 237 (2013), pp. 62–146. DOI: 10.1016/j.aim.2013.01.005.

[80] D. Kühn and D. Osthus. "Hamilton cycles in graphs and hypergraphs: an extremal perspective". In: *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. IV.* Kyung Moon Sa, Seoul, 2014, pp. 381–406.

[81] D. Kühn and D. Osthus. “Hamilton decompositions of regular expanders: applications”. In: *J. Combin. Theory Ser. B* 104 (2014), pp. 1–27. DOI: 10.1016/j.jctb.2013.10.006.

[82] F. T. Leighton. *Introduction to parallel algorithms and architectures*. Arrays, trees, hypercubes. Morgan Kaufmann, San Mateo, CA, 1992.

[83] C. McDiarmid, A. Scott, and P. Withers. “The component structure of dense random subgraphs of the hypercube”. In: *arXiv e-prints* (2018). arXiv: 1806.06433.

[84] S. Messuti, V. Rödl, and M. Schacht. “Packing minor-closed families of graphs into complete graphs”. In: *J. Combin. Theory Ser. B* 119 (2016), pp. 245–265. DOI: 10.1016/j.jctb.2016.03.003.

[85] R. Montgomery. “Fractional clique decompositions of dense graphs”. In: *Random Structures Algorithms* 54 (2017), pp. 779–796. DOI: 10.1002/rsa.20809.

[86] N. Pippenger and J. Spencer. “Asymptotic behavior of the chromatic index for hypergraphs”. In: *J. Combin. Theory Ser. A* 51 (2017), pp. 24–42. DOI: 10.1016/0097-3165(89)90074-5.

[87] L. Pósa. “Hamiltonian circuits in random graphs”. In: *Discrete Math.* 14 (1976), pp. 359–364. DOI: 10.1016/0012-365X(76)90068-6.

[88] D. Ray-Chaudhuri and R. Wilson. “The existence of resolvable designs”. In: *A survey of Combinatorial Theory (J.N. Srivastava, et. al., eds.)* North-Holland, Amsterdam, 1973.

[89] C. Savage. “A survey of combinatorial Gray codes”. In: *SIAM Rev.* 39 (1997), pp. 605–629. DOI: 10.1137/S0036144595295272.