# An introduction to the absorption method

Alp Müyesser[†]

October 31, 2025

### Abstract

Finding large transversals in Latin squares [Mon] and constructing designs [Kee14] constitute two landmark applications of the absorption method. In these notes, we will introduce the absorption method from scratch, beginning with the set-up of Dirac-type problems, and then build our way up to transversals and designs. These notes are work in progress.

## 1 What is the absorption method?

Suppose we are trying to construct an object satisfying a certain set of constraints. The absorption method gives us an angle of attack if the following two properties hold in our context.

(A) It is significantly easier to construct some object satisfying 99% of the constraints.

(B) There is a robust sense in which partial solutions can be *locally* modified to produce other partial solutions.

The above properties hold in a wide variety of contexts. For example, if the constraints come about in a structureless, uniform sort of way, a sensible approach towards (A) is simply *the probabilistic method*, i.e. considering a suitable notion of a randomly sampled object. If a random construction satisfies 100% of the constraints with positive probability[1], then of course there is no need for the absorption method. There are, however, several natural instances in which a random strategy gets us most of the way, and then gets stuck towards the end[2]. This is where (B) comes in to somehow turn the approximate solution into an exact one.

When is (B) satisfied? Let's start with a non-example. Consider the task of constructing a proper 2-vertex-colouring of a cycle. Here, there is no local flexibility whatsoever: the effect of changing the colour of a vertex necessarily has to propagate through the whole cycle, and hence has *global* effects[3]. As for a positive example, we will see several throughout these notes, beginning with the following beautiful argument of Erdős from 1986 [Rei89], which appears to be the first application of the absorption method (to my knowledge), though the wide applicability of the method was not crystallised until pioneering work of Rödl, Ruciński, and Szemerédi [RRS06] from 2006.

Recall that a *tournament* is a complete graph, where each edge $\{i, j\}$ is assigned a direction $i \to j$ or $j \to i$. A tournament is transitive if its vertices can be ordered as $(1, 2, \ldots, n)$ so that $i \to j$

---

[†]University of Oxford. email: alp.muyesser@new.ox.ac.uk

[1]For example, as is the case whilst constructing an $n$-vertex graph with no cliques or independent sets of size $2 \log_2 n$.

[2]As a first example, the reader may consider the puzzle of the *n-queens problem*, i.e. the task of placing $n$ mutually non-attacking queens on an $n \times n$ chessboard (see, for example, [SL22])

[3]As it should — whether we succeed or fail is of course entirely contingent on a global property: the parity of the number of the vertices.

if and only if $i < j$ (or equivalently, if a tournament is transitive if it contains no distinct $i, j, k$ with $i \to j \to k \to i$).

**Proposition 1.1** (Erdős, 1986). *For every $k$, there exists an $N$ so that for all $n > N$ such that $k$ divides $n$, the vertices of any $n$-vertex tournament can be partitioned into transitive sub-tournaments of size $k$.*

We need to first recall that any sufficiently large tournament contains a large transitive tournament. This can be seen from Ramsey's theorem by constructing an auxiliary red-blue edge colouring, but the following direct argument gives an even better bound.

**Observation 1.2.** *Suppose $n \geq 2^k$. Any $n$-vertex tournament contains a transitive sub-tournament on $k$ vertices.*

**Proof.** By induction, suppose the assertion holds for smaller values of $k$. Fix some transitive tournament of size $n \geq 2^k$. Take some arbitrary vertex $v$, and notice $v$ either has at least $n/2 \geq 2^{k-1}$ in-edges or out-edges, by the pigeonhole principle. Let's say $v$ has many in-edges. Invoke the inductive hypothesis on the in-neighbourhood to find a transitive sub-tournament there of size $k - 1$, and tack on $v$ to the end of this tournament to find one of size $k$. The other case is of course symmetric. $\square$

Given the above, how should we approach Proposition 1.1? Let us first check (A). Say we took $N = N(k)$ to be $100 \times 2^k$. Then, by greedily plucking transitive sub-tournaments of size $k$ one by one, we may saturate all but 1% of the original tournament with transitive sub-tournaments of the same size. So it feels indeed easier to satisfy 99% of the constraints.

How about (B)? Now suppose we took $N = 100 \times 2^{2k}$. Then we can repeat the same argument with transitive sub-tournaments of size $2k$ saturating 99% of the original tournament. Each transitive sub-tournament of size $2k$ splits into two disjoint ones of size $k$ in a rich variety of ways, $\binom{2k}{k}$ ways to be exact. So there indeed are many partial solutions that can be obtained from one another by making "local toggles".

It's not actually clear how we can use these properties (A) and (B) to our advantage to fully solve the problem. The reader may want to stop and think for a moment, before we reveal the brilliant strategy of Erdős.

**Proof of Proposition 1.1.** Given a large enough tournament $T$, let us first reserve $K := k2^{k-1}$ many disjoint transitive sub-tournaments $T_i$ ($i \in [K]$) of size $2k - 1$. Let us keep plucking disjoint transitive sub-tournaments of size $k$ in the remainder (this can be done as long as there are $2^k$ vertices remaining, but we will stop earlier). The leftover must stay at being multiples of $k$ (as $K$ also is a multiple of $k$), so at some point, the size of the leftover is exactly $K$, so we can choose to stop the process at that moment. Let us arbitrarily pair up the leftover vertices with the $T_i$. The key observation that concludes the proof is the following: for each leftover vertex $v$, $v + T_i$ can be split into two transitive sub-tournaments, as $v$ has at least $k - 1$ neighbours in $T_i$ with the same orientation of edges (in fact $k$, but $k - 1$ is what we need here). $\square$

The above proof consists of three steps that appear one way or another in most applications of the absorption method.

**The absorption method**

**Step 1.** (Preparation) Find some flexible structure that satisfies the constraints in a rich variety of ways.

**Step 2.** (99% solution) In the remainder, saturate most of the constraints.

**Step 3.** (Absorption) Show that the leftover from Step 2 combines with the flexible structure from Step 1 to produce a construction that satisfies all of the constraints.

A motivated reader might want to think of the quantitative aspects of Proposition 1.1. The Erdős argument certainly requires $N \gg 2^{2k-1}$, and one cannot do better than $N \gg 2^k$ without improving on Observation 1.2. Can we improve on the constant factor in the exponent with a different argument and replace it with $1 + o(1)$? See [BS21] for an answer.

## 2 Perfect matchings in (hyper-)graphs and a digression on the switching method

Both the study of transversals in Latin squares [Mon] and the constructions of designs [Kee14] turn out to be expressible in the language of *hypergraph perfect matchings*. In this section, we give an overview of hypergraph perfect matchings, starting with the graph case, where simple "switching-methods" give satisfying answers, and the hypergraph case, where the said methods fail. We also discuss how the absorption method can in some sense be thought as a more robust implementation of the switching method, relying only on "local variability" to engineer some "global variability".

A *spanning subgraph* of a graph $G$ is a subgraph that has the same vertex set as $G$ and an edge set that is contained in $E(G)$. A decent chunk of extremal combinatorics seeks to characterise (hyper-)graphs that contain a certain kind of spanning subgraph. For example, a graph contains a tree as a spanning subgraph if and only if the graph is connected.

As a more substantial example, we have Hall's marriage theorem. Recall that if $S \subseteq V(G)$, the neighbourhood $N(S)$ denotes elements of $V(G)$ that are connected to some vertex in $S$.

**Theorem 2.1** (Hall's marriage theorem). *A bipartite graph $G = (A, B)$ with $|A| = |B|$ has a perfect matching if and only if the following condition holds.*

(Hall's condition.) *For every $A' \subseteq A$, $|N(A')| \geq |A'|$.*

A direct corollary of Hall's marriage theorem is that it is always possible to give a short certificate that a bipartite graph does **not** contain a perfect matching[4]: one simply points to the set $S$ with $|N(S)| < |S|$. This is very rare phenomenon in the study of spanning structures. For example, no analogous statement holds for spanning paths, or higher dimensional perfect matchings, subject to mainstream hypotheses in computer science.

To formalise this a bit more, a *tripartite hypergraph $H = (A, B, C)$* is a collection of triples $E(H) := (a, b, c) \in A \times B \times C$ which we view as being the (hyper-)edges of $H$, and $A$, $B$ and $C$ are assumed to be disjoint. The definition of a perfect matching hasn't changed, it still is a collection of vertex-disjoint edges that together cover the vertex-set $A \cup B \cup C$. We call a tripartite graph *balanced* if $|A| = |B| = |C|$, an obvious necessary condition for the existence of a perfect matching.

Existence of perfect matchings in tripartite hypergraphs is one of Karp's original 21 NP-complete problems. What this means for the working combinatorialist is that there almost certainly is no nice characterisation of tripartite hypergraphs with perfect matchings, and moreover, in general, we expect that it is hard to prove (or give a certificate) that a given tripartite hypergraph does **not** contain a perfect matching. That said, many problems in combinatorics can be expressed as a hypergraph matching problem including the existence of transversals in Latin squares and the existence of designs, the two subjects that we are trying to build our way into. So, we certainly wish for a better understanding of hypergraph matchings, no matter how weak.

A widely useful criterion for when tripartite hypergraphs contain perfect matchings is given in the work of Aharoni–Haxell, and these results are interesting not least because they exemplify how topological methods can come into play for problems that seem entirely combinatorial in nature. Unfortunately, these methods and criteria have thus far not been very applicable in the study of transversals in Latin squares or designs, hence we will not say more about them here.

---

[4]In complexity-theoretic terms, the language of bipartite graphs with perfect matchings is in the class co-NP.

Let us set our targets a bit lower and consider the following (essentially trivial) corollary of Hall's marriage theorem.

**Proposition 2.2.** *If $G = (A, B)$ is a d-regular bipartite graph (each vertex is incident on exactly d-edges), then G has a perfect matching.*

**Proof.** Observe $|A| = |B|$ as $e(G)$ (number of edges of $G$) is equal to both $|A| \cdot d$ and $|B| \cdot d$. Also, for all $A' \subseteq A$, $e(A', N(A'))$ (number of edges between $A'$ and its neighbourhood, $N(A')$) is equal to $d \cdot |A'|$, and is at most $|N(A')| \cdot d$, so $|N(A')| \geq |A'|$. $\qquad\square$

Most hypergraphs we consider here will be regular, so it would have been convenient to have any high-dimensional analogue of Proposition 2.2. Unfortunately, such a result cannot exist, even when we restrict our attention to *linear* hypergraphs, meaning those hypergraphs where any two vertices are contained in at most one common edge. In fact, such a result fails even when we restrict our attention to $n$-regular linear hypergraphs $H = (A, B, C)$ where $|A| = |B| = |C| = n$, which can be interpreted as the counter-intuitive assertion that there is no analogue of $K_{n,n}$ having a perfect matching for linear tripartite hypergraphs.

**Example 2.3.** *Consider a tripartite hypergraph $H = (A, B, C)$ where $A, B, C$ are disjoint copies of $\mathbb{Z}/n\mathbb{Z}$ where n is an even integer and $(a, b, c)$ is an edge if and only if $a + b + c = 0 \pmod n$.*

**Proof.** If $H$ had a perfect matching, summing over all edges of the matching we'd obtain $3 \cdot \sum_{x \in \mathbb{Z}/n\mathbb{Z}} x = 0 \pmod n$, but $3 \cdot \sum_{x \in \mathbb{Z}/n\mathbb{Z}} x \equiv 3 \cdot (n/2) \equiv n/2 \not\equiv 0 \pmod n$, assuming $n$ is even. $\qquad\square$

Considering disjoint unions of the above construction when $n = 2$, we can see that in general, a regular linear tripartite hypergraph can be quite far away from containing a perfect matching, the largest matching could miss half the vertices! We would expect though the situation would get less bad if the degrees of the hypergraph were larger. For example, any $n$-vertex $d$-regular graph (not necessarily bipartite) contains a matching covering all but $O(n/d)$ vertices, as can be deduced from theorems of Vizing, or Tutte (and this is optimal, as the graph could consist of disjoint unions of $O(n/d)$ odd-sized subgraphs).

A fundamental open problem of Alon–Kim–Spencer asks about the size of the largest matching in a general $d$-regular $n$-vertex linear hypergraph (of uniformity 3). Alon–Kim–Spencer showed that there is a matching covering all but $nd^{-1/2+o(1)}$ vertices (with techniques related to the Rödl nibble that we will discuss later on), but, in principle, it may be possible to find matchings covering all but $O(n/d)$ vertices. Even when $d = n$, proving such a result was a topic of extensive study, culminating in Montgomery's tour de force work [Mon] proving the following.

**Theorem 2.4** (Montgomery, 2023+). *Any n-regular linear balanced tripartite hypergraph on 3n vertices contains a matching with $n - 1$ edges.*

A striking conjecture of Ryser is that if $n$ is odd, $n - 1$ can be replaced with $n$ in the above, meaning constructions like Example 2.3 can only exist for even values of $n$.

We remark that linear tripartite hypergraphs $H = (A, B, C)$ with $n = |A| = |B| = |C|$ that are $n$-regular are often presented as *Latin squares*. An order $n$ Latin square is an $n \times n$ array filled with $n$ symbols so that no symbol repeats in any row or any column. Thinking of $(A, B, C)$ as (rows, columns, symbols), and $s$ being the symbol in row $r$ and column $c$ as $(r, c, s)$ being an edge, we see that Latin squares and $n$-regular linear tripartite hypergraphs on $3n$ vertices are the same objects. Perfect matchings are often called *transversals* in the Latin square language. We motivated the study of perfect matchings through the graph theoretic lens in the above, but there are also strong design-theoretic motivations to study transversals in Latin squares, as

they have applications in error-correcting codes and the design of actual scientific experiments where the goal is to minimise unwanted correlations.

## 2.1  A digression on the switching method

We provide the short proof of Hall's marriage theorem, as it allows us to expose a close cousin of the absorption method, namely, the switching method.

**Proof of Theorem 2.1.** Let $M$ be a matching of $G$ (not necessarily perfect), and let $F := E(G) \setminus M$ be the unmatched edges. An $M$-alternating path is a path of $G$ whose edges alternately come from $F$ and $M$. An $M$-alternating path is an $M$-augmenting path if it starts and ends with vertices not matched by $M$. Given an augmenting path, we may swap the unmatched edges with matched edges, and find a matching $M'$ with $|M'| > |M|$. So (the non-trivial direction of) Hall's marriage theorem reduces to showing that any non-perfect matching $M$ in a graph $G$ with Hall's condition has an $M$-augmenting path.

So let $M$ be a matching of $G$ and suppose $v \in A$ is unmatched. Consider the set of all vertices $R_v$ that are reachable by an $M$-alternating path starting from $v$, and note that if $R_v$ contains an unmatched vertex $w \in B$, we're done. Consider the sets $R_v \cap A$ and $R_v \cap B$. The key observation is that unless $R_v \cap B$ contains an unmatched vertex from $B$, it must be that $R_v \cap B$ and $R_v \cap A$ are in bijective correspondence with the pairing coming from $M$ (the best way to see this is to draw a picture and track the iterative growth of $R_v$ as we allow longer and longer alternating paths). Also, $R_v \cap B$ contains $N(v)$, as well as $N(R_v \cap A)$, as any $G$-neighbour of $v$ or $R_v \cap A$ is certainly reachable by a $M$-alternating path from $v$. We then find a set of size $|(R_v \cap A) \cup \{v\}|$ whose neighbourhood has size $|(R_v \cap A)|$, contradicting Hall's condition, as desired. □

The above proof of Hall's marriage theorem belongs in a wider class of proof strategies, which can be summarised as follows.

> **The switching method**
> **Step 1.** Find some object maximal with respect to some parameter which we are trying to optimize.
> **Step 2.** Show that unless the object satisfies all the constraints, it can be improved by performing a sequence of modifications.

In the context of structural graph theory, the switching method is studied under the guise of "Kempe changes" (for example, in the proof of Vizing's theorem on edge-colourings or the 5-colour-theorem), and in the context of extremal graph theory, the switching method often manifests itself as the study of "Posá rotations" (we will see an illustration in the next section).

The third step in the absorption method is often realised by performing a sequence of *local* modifications, so the absorption method can really be thought of as a special case of the switching method. The proof of Hall's marriage theorem features a basic application of the switching method that requires no preparation, nor the 99% solution step, mainly because there is a tractable way in which we can reason about the space of *global* modifications we can perform. Indeed, neither the length of the augmenting path nor the size of the matching $M$ being augmented had any role in the argument.

The absorption method is necessitated by the fact that in many contexts, there is no known tractable way of reasoning about global modifications[5], yet there is a way of reasoning about local modifications.

---

[5] And moreover, there are compelling philosophical reasons for why there ought to not be a way to think about such global modifications in a tractable way, with evidence coming from computational complexity theory.

The selling point of the absorption method is that if we play our cards right in the preparation step, and if we furthermore facilitate what we have to do in the "modification stage" of the proof by dealing with 99% of the constraints already, then we can also find a way to deal with the remaining 1%. "Playing our cards right" often entails chaining up local objects that display local variability together to produce a global object that displays global variability. We already saw a basic manifestation of this phenomenon in the proof of Proposition 1.1, and we will now turn our attention to slightly more involved examples.

# 3 Dirac-type theorems

We want to build our way up to understanding aspects of Montgomery's theorem (Theorem 2.4), but to get there, we will first build more intuition on the absorption method. The absorption method was first codified by Rödl, Ruciński, and Szemerédi [RRS06] to generalise a classical theorem of Dirac to hypergraphs, and the set-up of these kinds of problems serve as a great learning ground. In particular, here, we will develop the idea of a *reservoir* and also a *cover-down*, which will be quite important later on.

**Theorem 3.1** (Dirac's theorem, 1950s)**.** *If G is a n-vertex graph of minimum degree at least $n/2$, then G has a spanning (Hamilton) cycle, meaning a cycle that hits all vertices of the graph.*

Graphs with minimum degree less than $n/2$ need not even be connected, so Dirac's theorem is sharp, as far as sufficient conditions for Hamiltonicity in terms of the minimum degree go.

The proof of Dirac's theorem is fairly simple — just like the proof of Hall's marriage theorem, it makes use of switchings, adapted to the context of path-like structures, where they are often called *rotations*.

**Proof of Dirac's theorem.** As the minimum degree is at least $n/2$, and two vertices have at least one common neighbour, in particular, the graph is connected. We will show that for any maximal length path $P$, there is a cycle on the vertex set $V(P)$. Connectivity then implies the existence of a spanning cycle (why?).

Take a maximal length path $P$, then, and suppose $u$ and $v$ are its endpoints. All $\geq n/2$ neighbours of $u$ and $v$ have to be on the path $P$ by maximality, and if $u \sim v$, we are happy. Otherwise, we do a rotation argument as follows. By the pigeonhole principle, there exist consecutive vertices of $P$, $x$ and $y$ (where $x$ is the vertex closer to $u$) such that $x \sim v$ and $y \sim u$. Traversing the cycle as $u \to y$, following $P$ until $v$, $v \to x$ and following $P$ until $u$ again closes a cycle on the same vertex set, as desired. $\square$

There is a rich literature on higher-dimensional generalisations of Dirac's theorem in the literature, in terms of obtaining best-possible minimum degree conditions that imply the existence of a certain spanning subgraph. We will not attempt to give a survey of results, but a brief survey of the ideas that were generated as a result of this research programme.

As appealing as the proof of Dirac's theorem is, it unfortunately does not seem to generalise to higher dimensions, or at least, nobody has figured out a way to make such an argument work yet. Up until the absorption method became prevalent through works of Rödl, Ruciński and Szemerédi [RRS06], the mainstream method for proving higher-dimensional generalisations of Dirac's theorem were the regularity/blow-up method. This approach was pioneered by Komlós, Sarközy, and Szemerédi, and it entails working with a partition guaranteed by the Szemerédi regularity lemma, embedding most of the desired structure via a randomised algorithm, and then showing that there is a way to finish things off by reducing the remaining task to an application of Hall's marriage theorem. There exist hypergraph generalisations of the Szemerédi regularity lemma, and extensions of the blow-up method by Keevash, but such an approach quickly becomes prohibitively technical, and moreover, it does not generalise to the set-up of transversals or designs (as the corresponding hypergraphs here are fairly sparse, making Szemerédi-type regularity lemmas inapplicable), so we will not say more about the blow-up methods here.

The development of the absorption machinery has led to a fairly satisfying theory of sufficient conditions for Hamiltonicity in dense hypergraphs, see for example the recent work of Lang and Sanhueza-Matamala [LSM24]. This theory, very roughly speaking, goes as follows. We know that any Hamiltonian graph needs to be (1) connected, and depending on the parity of

the number of vertices, (2) the existence of a perfect matching[6], and (3) an odd-sized cycle is also forced. It turns out that any graph satisfying (1-3) in a very robust sense has to be Hamiltonian. The robustness property is formalised by asserting that almost all small-sized subsets inherit the properties (1-3), which often holds in dense and quasirandom settings. A version of this theory adapted to hypergraphs can concisely capture many high-dimensional extensions of Dirac's theorem in the literature. Although this theory is powerful, it does not tell us much about how to ensure the properties (1-3), which in some cases yield fairly difficult problems in extremal hypergraph theory, many of which are still unsolved.

In the simple context of Dirac's theorem though, the three properties are fairly simple to establish, allowing us to give a straightforward absorption proof of the following asymptotic relaxation of Dirac's theorem.

**Theorem 3.2.** *For every $\varepsilon > 0$, for $n$ sufficiently large, any $n$-vertex graph $G$ with $\delta(G) \geq (1+\varepsilon)n/2$ has a Hamilton cycle.*

Of course this is weaker than Dirac's theorem, and the proof will be more convoluted than the simple rotation argument we just gave. Our motivation is as follows:

- Minor changes to the proof actually establish the sufficiency of (1-3) discussed above[7], which significantly extends the minimum degree set-up considered in Dirac's theorem.
- Often, one can find a way to turn asymptotic results as Theorem 3.2 into exact result, by performing what's called a *stability analysis*. In the set-up of Dirac's theorem, this in not particularly difficult, though in higher-dimensions this can get a bit tricky, though such a strategy was successfully pursued in many contexts.
- Recall that our real motivation is to learn the absorption method better, so that we can better understand transversals and designs, and the proof will demonstrate a transparent application of *the reservoir method*, which we will return to on several occasions.

To build some intuition for the proof, let us go back to (A) and (B), and see in what sense they hold in the context of Theorem 3.2, starting with (A).

## 3.1   99% **solution**

We have the following easy proposition.

**Proposition 3.3.** *If $G = (A, B)$ is a bipartite graph with $|A| = |B| = n$ of minimum degree at least $n/2$, then $G$ has a perfect matching.*

**Proof.** This can be seen by checking Hall's condition, or directly by considering a largest matching as follows. Supposing there exists vertices $u$, $v$ that are unmatched, their entire neighbourhood must be contained in the vertex set of the largest matching. By the pigeonhole principle, there must exist an edge of the matching that $u$ and $v$ are both adjacent to. This corresponds to an augmenting path on 3-edges, so performing a switch guarantees a larger matching, a contradiction. □

Our proof will rely on relatively small subsets of $G$ containing perfect matchings if $\delta(G) \geq (1+\varepsilon)n/2$.

---

[6]We could also work with fractional perfect matchings here, which are weight functions on the edge set so that the sum of weights along the edges incident on any vertex is exactly 1.

[7]One should note that for higher-dimensional generalisations of this theory, one needs to work harder, in particular our argument does not generalise straightforwardly to higher dimensions, but the argument should still give a good introduction to start reading the work of Lang and Sanhueza-Matamala [LSM24]

**Proposition 3.4.** *Let $v$ have $(1 + \varepsilon)|G|/2$ neighbours in $G$, and let $S \subseteq V(G)$ be a subset of size $C$, sampled uniformly at random. Then, with probability $1 - \exp(-\Omega(\varepsilon C))$, $v$ has at least $(1 + \varepsilon/2)|S|/2$ neighbours in $S$.*

**Proof.** This is essentially direct by Chernoff's bound, though a slight nuance is that we consider a *hypergeometric distribution*, meaning $S$ is sampled not independently, but subject to having a certain size. As one might expect, tail-bounds do not change much in this set-up, as one can formally show by applying Azuma's inequality or related tools. For example, a result of McDiarmid says that if a random variable is a function of a randomly sampled subset of a ground set of a certain size, and the value of the random variable changes only a little bit if an element of the input subset is interchanged, then we get strong concentration bounds as claimed above, see [LW17, Lemma 6.1] for a formal statement and the derivation from the classical Azuma martingale inequality. □

Using the above and a union bound, we get the following corollary.

**Corollary 3.5.** *Let $A, B$ be disjoint and random subsets of $V(G)$ each of size $C$ where $\delta(G) \geq (1 + \varepsilon)n/2$. With probability $\geq 1 - 2C \exp(-\Omega(\varepsilon C))$, there exists a perfect matching between $A$ and $B$.*

Let us imagine $C = (\log n)^2$, and that $C$ divides $n$. In this case, given a $n$-vertex $G$ with $\delta(G) \geq (1 + \varepsilon)n/2$, we may partition $V(G)$ randomly into sets of size $C$ as $A_1, A_2, \ldots, A_{n/C}$, and invoke Corollary 3.5 between every pair $G' = (A_i, A_{i+1})$ to find a perfect matching between the specified sets (the choice of $C = (\log n)^2$ allows us to comfortably take a union bound over all $(A_i, A_{i+1})$ (check this!)). Taking a union of all the perfect matchings, we have managed to cover $G$ with $\log^2 n$ paths, or in another perspective, we found a Hamilton cycle, minus $(\log n)^2$ edges, which feels like good enough progress to say that (A) holds in the context of Theorem 3.2. We can go even further, and chain up these paths into a cycle that is *almost* spanning. First, we need a simple lemma stating we can set aside a random set through which all pairs of vertices are well-connected.

**Proposition 3.6** (Reservoir exists)**.** *Let $G$ have minimum degree $\geq (1 + \varepsilon)|G|/2$, and let $S \subseteq V(G)$ be a subset of size $C$, sampled uniformly at random. Then, with probability $1 - |G|^2 \exp(-\Omega(\varepsilon C))$, the following property holds:*

    **P** *Every $u, v \in V(G)$ have at least $(\varepsilon/4)|S|$ common neighbours in $S$.*

**Proof.** Observe that any two vertices $u, v \in V(G)$ has at least $(\varepsilon/2)|G|$ common neighbours in $G$. The assertion about the random set $S$ picking up an appropriate number of common neighbours for each pair of vertices then comes from a concentration and a union bound argument, same as the one we just saw before. □

We will think of $S$ as a "reservoir" of vertices that we can use to stitch up the paths we obtain in the argument spelled out after Corollary 3.5. With a bit of foresight, we prove a lemma that also takes as input two vertices $u, v$ that are to be endpoints of a long path covering all vertices outside of $S$. As the reader might guess, we remark that the specific choice of the parameters is not particularly important, only the relationship between the parameters, and that they are at least polylogarithmic (so that a union-bound can be taken).

**Lemma 3.7** (Cover-down step)**.** *Let $G$ have minimum degree $\geq (1 + \varepsilon)n/2$ where $n := |G|$ and let $S \subseteq V(G)$ be of size $C := n^{1/100}$ satisfying **P**. With probability $1 - o_n(1)$, the following holds. For any $u, v \in V(G) \setminus S$, there exists a path $P$ with endpoints $\{u, v\}$, and $V(G) \setminus S \subseteq V(P)$.*

**Proof.** We know $S$ satisfies **P**, so every pair of points have at least $\varepsilon n^{1/100}/4$ common neighbours in $S$)

Let us suppose for simplicity that $n = n^{1/100} + (\log n)^2 \cdot t$ for some integer $t \sim n/(\log n)^2$.

Partition $V(G) \setminus S$ further into disjoint random sets $A_1, \ldots, A_t$ of size $(\log n)^2$ each (noting the graph restricted to the complement of $S$ has minimum degree $(1 + \varepsilon/2)|G|/2$. , and with high probability Corollary 3.5 holds for each pair $(A, B) = (A_i, A_{i+1})$ simultaneously by a union bound.

Fix now an outcome of the random sampling satisfying these properties. We can then find a perfect matching in each $G = (A_i, A_{i+1})$ $(1 \le i \le t - 1)$, yielding a collection of $(\log n)^2$ vertex-disjoint paths partitioning $V(G) \setminus S$ (henceforth called spanning path forest). At the cost of increasing the number of paths to $(\log n)^2 + 2$, we can ensure that $u$ and $v$ are both endpoints of a *distinct* path in the spanning path forest, let's call it $\mathcal{P}$.

Let $\{x_i, y_i\}_{i \in I}$ be a pairing of the endpoints of $\mathcal{P}$ except for $u$ and $v$, so that if we "merge" each of the endpoints $\{x_i, y_i\}$, the resulting object would be a path with endpoints $u$ and $v$ (convince yourself such a pairing exists!). Now, we claim that there exists a collection of vertex disjoint paths $\{P_i\}_{i \in I}$, each on two edges, with endpoints $\{x_i, y_i\}_{i \in I}$ and the unique internal vertex in $S$. That $\{P_i\}_{i \in I}$ exists follows from the property of $S$: for each $\{x_i, y_i\}$, we can find a common neighbour that wasn't previously used on a "connecting path", as there are $\le (\log n)^2$ connecting paths and $\ge \varepsilon n^{1/100}/4$ common neighbours, one can always find the next required connecting path. □

The statement of Lemma 3.7 contains a fairly important "cover-down" idea that will be reoccur both whilst studying transversals, and existence of designs. Applying the Lemma with $u \sim v$ being an arbitrary edge of $G$, we see that we can find a cycle that covers all but a few vertices of $G$, and moreover, the leftover vertices are restrained to a small subset that was pre-defined. We will later see a bootstrapping variation on this idea called "iterative absorption".

## 3.2 Absorption

How about (B)? This is where the odd-cycles come in. In the setting of Theorem 3.2 we indeed have many odd-cycles, even triangles. Indeed, any vertex $x$ must be contained in several triangles if the minimum degree is slightly more than half. We call a subgraph $T \subseteq G$ with $u, v, x \in V(T)$ a $(\{u, v\}, x)$-absorber if $T$ contains a spanning path, and $T \setminus \{x\}$ contains a spanning path, and both spanning paths have $\{u, v\}$ as its endpoints. Observe that for any triangle $T$ with vertices $u, v, x$, is a $(\{u, v\}, x)$-absorber. $(\{u, v\}, x)$-absorbers being plentiful across $G$ for any $x \in V(G)$ is the key "local variability" property we will use in the proof. We recommend the reader the challenge of constructing more delicate $(\{u, v\}, x)$ absorbers that contain no triangles, in fact, no short cycles at all.

> **Exercise 3.8.** *Show that for every $k$ there exists a graph $T$ and vertices $u, v, x \in V(T)$ so that $T$ is a $(\{u, v\}, x)$-absorber, $T$ has at most $O(k^2)$ vertices, and the shortest cycle in $T$ has length at least $k$.*

The motivation for the above exercise is to justify our assertion that odd-cycles being abundant and connectivity being inherited by typical small induced subgraphs is all that is necessary to find Hamilton cycles (in addition to the perfect matchings).

Putting the individual local absorbers together, we can obtain a global absorber for a large reservoir. In general, chaining up local absorbers to obtain such a global property won't be so easy, but in the set-up of Dirac's theorem, it is fairly transparent how this takes place.

**Definition 3.9.** We call a subgraph $T \subseteq G$ with $u, v \in V(T)$, $S \subseteq V(T)$ a $(\{u, v\}, S)$-absorber where $S \subseteq V(G)$ if $T \setminus S'$ contains a spanning path for each $S' \subseteq S$, and each such spanning

path has $\{u, v\}$ as its endpoints.

**Lemma 3.10.** *Let G have minimum degree $\geq (1 + \varepsilon)n/2$ where $n := |G|$ and let $S \subseteq V(G)$ be of size $\leq \gamma n$ where $\gamma \leq \varepsilon/100$. Then, there exists a subgraph T that is a $(\{u, v\}, S)$-absorber for some $u, v \in V(G)$ where $|T| \leq 5|S|$.*

**Proof.** First find a collection of disjoint triangles where each triangle contains a unique element of $S$. Then connect the appropriate pairs of vertices with paths of length two so that the shape in Figure 1 emerges. The simple details are left to the reader.
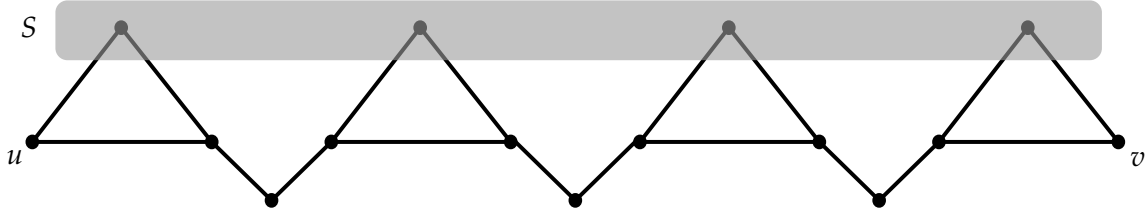


Figure 1: The construction of the $(\{u, v\}, S)$-absorber. The reader should check that such an object exists under the specified minimum-degree condition whenever $S$ is a subset of size at most $(\varepsilon/100)n$.

The reader should also check that for each subset $S' \subseteq S$ in the above, there exists the desired spanning path "absorbing" $S \setminus S'$. □

## 3.3 Putting things together

**Proof of Theorem 3.2.** Set aside a set $S$ of size $C := n^{1/100}$ with the property **P**, the existence of $S$ is guaranteed by Proposition 3.6. By Lemma 3.10, there exists a subgraph $T$ of size at most $5C$ that is a $(\{u, v\}, S)$-absorber for some $u, v$. Applying Lemma 3.7 with the graph $G$ with vertices of $T \setminus \{u, v\}$ deleted (observing this does not meaningfully change the minimum degree of $G$), and with $\{u, v\}$, we find a path $P$ with endpoints $\{u, v\}$, fully saturating $G \setminus T \setminus S$, and using some vertices of $S$. The unused vertices of $S$, meaning $S' := S \setminus V(P)$, combine with $T$ to find a $\{u, v\}$ path $P'$. $P' \cup P$ is then the desired Hamilton path. □

## 3.4 Historical aside

The reservoir/cover-down style of absorption is not how the absorption method was originally popularised by Rödl, Ruciński, and Szemerédi (RRS) [RRS06]. They had a different notion of an "absorbing path", which can absorb vertices coming from anywhere along the graph, as opposed to just a predefined reservoir, as long as the set of vertices to be absorbed is small enough, compared to the size of the absorber. On the other hand, in the above construction, our absorber could absorb any subset of vertices, although they had to live in the reservoir. The reservoir perspective will be more useful later on, but obtaining a RRS-style absorber also isn't too difficult in the setting of Dirac's theorem.

**Definition 3.11.** Given an ambient graph $G$, we call a subpath $P$ to be a $\gamma$-absorber if for all subsets $S$ of $\leq \gamma|G|$ vertices, $S \cup P$ has a spanning path with the same endpoints as those of $P$.

> **Exercise 3.12.** *Let G be a graph of minimum degree $(1 + \varepsilon)|G|/2$. Show that there exists a $\gamma := \gamma(\varepsilon)$ so that G contains a $\gamma$-absorber path P of size at most $\varepsilon|G|/2$.*
> **Hint:** *Sample a set of $f(\varepsilon)|G|$ random edges, check that for every vertex, some edges that form triangles with that vertex were sampled. Delete the edges around vertices of degree $> 2$, and chain up the remainder into a path.*

# 4   Concentration of measure and the Rödl nibble

We are trying to build towards understanding Latin squares that have full transversals (of size $n$) and near transversals (of size $n-1$). But before, we need to first be able to demonstrate that all Latin squares have transversals of size $0.99n$, which checks (A) and starts to make the absorption method potentially applicable. The most basic form of such a statement would be as follows. Recall that "Let $n^{-1} \ll \gamma$" in the beginning of a statement means for any $\gamma > 0$, there exists $n^{-1}$ sufficiently small so that the following statement holds.

**Proposition 4.1.** *Let $n^{-1} \ll \varepsilon$. Any 3n-vertex n-regular tripartite linear hypergraph (i.e. Latin square) has a perfect matching (i.e. transversal) covering all but $\varepsilon n$ vertices.*

We will prove results like the above via the *semi-random method*, or the *Rödl nibble*. The rough idea is we'll see (once we've built some intuition) that if we build a matching by a *random greedy strategy*, i.e. choosing an edge at random among all options that are disjoint with all previously selected edges, then we should get a desirable outcome. Analysing the random greedy strategy is however often tricky (though in many cases, including the above proposition, have been done, and isn't actually so technical with the above choice of parameters). The idea of the Rödl nibble is to group the random choices into big batches to ease the analysis. Think choosing $\varepsilon n$ random edges in the above (the bite). Then, do this about $1/\varepsilon n$ times and hope everything goes well (the nibbling). The gain is that the number of randomness stages to analyse has been reduced from $n$ to $1/\varepsilon = O(1)$ — so the errors coming from the deviations from the mean have much less time to accumulate.

This is an immensely powerful trick and finds applications quite widely whenever one wishes to analyse a random greedy algorithm but is willing to compromise a little, see for example the work of Ford, Green, Konyagin, Maynard, Tao on long gaps between primes, or the work of Campos, Jennsen, Michelen, and Sahasrabudhe on "amorphous" sphere packings in high dimensions.

Something stronger about nearly-regular hypergraphs is true and follows from the Rödl nibble, and the *hierarchy* of parameters here is useful to internalise. A hypergraph is $(\delta \pm \gamma)n$-regular if all degrees are in the range $[(\delta - \gamma)n, (\delta + \gamma)n]$.

**Proposition 4.2.** *Let $n^{-1} \ll \gamma \ll \delta, \varepsilon$. Any 3n-vertex $(\delta \pm \gamma)n$-regular tripartite linear hypergraph (can be viewed as a partially filled Latin square) has a perfect matching covering all but $\gamma n$ vertices.*

The important thing in the above is that $\gamma \ll \delta$, so that the degrees are actually concentrated in a small space, and $\gamma \ll \varepsilon$ — if we are hoping the complete $(1 - \varepsilon)$ fraction of the target process, the "unpredictability" in the degrees should better be much better constrained then a $\gamma n$ space for $\gamma \ll \varepsilon$.

In reality, we will need to prove both a quantitative and qualitative strengthening of such a result so the absorption method can meaningfully be applied.

Firstly, we will need to be able to find transversals of size $(1 - n^{-\varepsilon})n$ for some small constant $\varepsilon > 0$ — this is because it will turn out that our definition of "local" switches will need to allow for chains of switches up to $\log n$ coordinates, something that hasn't appeared in the basic applications we have seen thus far (where switches have always been of order $O(1)$). This in turns translates to being able to deal with leftovers from the 99% part only when the leftover is $\ll n/\text{polylog}(n)$, which is the case when it is $\ll n^{1-\varepsilon}$. Thankfully, the probabilistic method does just give stronger bounds such as these, so we won't need to work extra hard for this.

Secondly, we will need to deal with a certain *hierarchy issue*. Recall the basic premise of the absorption method: we first do some preparation to build some flexible structure $\mathcal{F}$, do 99% of the task, and show that the leftover combines with $\mathcal{F}$ to produce a full solution. In both

applications of the method we've seen, it was important that the leftover was comparable, or a fair bit smaller than the flexible structure $\mathcal{F}$, and this is a fairly general phenomenon. That is, if we are working in a $n$-vertex graph, find a "flexible absorber" $\mathcal{F}$ of size $an$, $\mathcal{F}$ is only able to "absorb" sets of size up to $bn$ for $b \ll a$.

> The reader might want to stop and think for a moment about the hole in the strategy of applying something like Proposition 4.2 *after* a flexible absorber is set aside in the beginning of the argument.

The issue is, when we are setting aside $\mathcal{F}$ of size $an$, assuming a worst-case condition, $\gamma$ in Proposition 4.2 will become at least as large as $a$, so then $\varepsilon$ will become $\varepsilon \gg a$. So if we naively set an absorber aside and hope everything will be okay, we will be disappointed in the end as we will end up with a leftover much larger in size than the absorber we set aside, rendering the absorber effectively useless.

How we deal with this issue in the set-up of transversals vs. designs will be quite different – but in both cases, we need a *regularity-boost* idea. Something that we do or observe after the absorber has been set aside, to tame the hierarchy of our parameters once again. In the setting of the Latin square hypergraph, this will boil down to noticing some *quasirandomness* properties random subgraphs of Latin squares have, which using the standard theory of quasirandomness (here we opt for that of Thomason (i.e. jumbledness)) will allow us to say that even in the worst case, setting aside an absorber won't damage the hierarchy all that much.

Before we can analyse the nibble and prove an appropriate strengthening of Proposition 4.2 compatible with our absorption strategy, we need to first review the concentration of measure toolkit.

## 4.1   Toolkit for concentration

The simplest one of all is Markov's inequality that says for a non-negative random variable $X$, we have $\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$ where $a > 0$ is arbitrary. This one really is quite transparent: for example, it says probability that $X$ is twice its expectation is at most $1/2$. As this considers only the expectation and nothing else, we call it a first-moment bound. If we are in a position where we can reason about $\mathbb{E}[(X - \mathbb{E}[X])^2]$, then we can apply Markov to the random variable $(X - \mathbb{E}[X])^2$, and get more powerful bounds in terms of the variance, a second-moment bound (also called Chebyschev's inequality). If our random variable breaks up into a sum of nice little independent chunks $X_i$ with $P(X_i = 1) = p_i$ and $P(X_i = 0) = 1 - p_i$, computing the second-moment is actually very easy. But why stop there, at that point we can compute for any $t$, $\mathbb{E}[e^{tX_i}]$, the moment generating function, and apply Markov to that. Point is that if $\sum X_i = X$, $e^{tX}$ can be expressed as a product $e^{tX_i}$, and the expectation of that product splits into the expectation of the individual pieces if the $X_i$ are independent. So we have:

**Lemma 4.3** (Chernoff bound). *Let $X := \sum_{i=1}^{m} X_i$ where $(X_i)_{i \in [m]}$ is a sequence of independent indicator random variables with $\mathbb{P}(X_i = 1) = p_i$. Then, for any $0 < \gamma < 1$, we have that*

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq \gamma \mathbb{E}[X]) \leq 2e^{-\mathbb{E}[X]\gamma^2/3}$$

.

We use the following corollary of Chernoff's bound often: that if $R$ is a $p$-random subset of an $n$-element set ($p$-random means each element sampled independently with probability $p$), then with high probability we have that $|pn - |R|| \leq \log n \sqrt{n}$.

Now onto more complicated random variables with dependencies. We have a product probability space $\Omega = \prod_{i=1}^{n} \Omega_i$ (this simply means we have $n$ independent trials) and $X \colon \Omega \to \mathbb{R}$

is a random variable. It's important to digest what's being said here, we are saying $X$ **is some random variable that's determined by $n$ independent trials**, but that's all we're assuming.

$X$ is called $C$-**Lipschitz** if $|X(\omega) - X(\omega')| \leq C$ whenever $\omega$ and $\omega'$ differ in at most 1-coordinate. We will refer to the following standard bound as McDiarmid's inequality.

**Lemma 4.4** (McDiarmid's inequality). *Let $X$ be $C$-Lipschitz random variable on a product probability space with $n$ coordinates. Then, for any $t > 0$,*

$$\mathbb{P}(|X - \mathbb{E}(X)| > \gamma \mathbb{E}[X]) \leq 2e^{\frac{-\gamma^2 \mathbb{E}[X]^2}{nC^2}}.$$

McDiarmid applies more generally than Chernoff, but if $\mathbb{E}[X] = o(n)$, (where $n$ is the number of coordinates in our product probability space), it starts to perform much worse quantitatively. It can be proved in a similar way to Chernoff's bound. The dependency on $n$ can be a deal-breaker in many situations we want to apply McDiarmid's inequality. In such situations, Talagrand's inequality comes to the rescue, whose proof is non-trivial, and makes crucial use of a "certifiability" assumption.

**Lemma 4.5** (Talagrand's inequality). *Let $X$ be a non-negative r.v. determined by $n$ independent trials, and satisfying the following for some $r, c > 0$.*

1. *$X$ is $C$-Lipschitz.*
2. *For any $s$, and sample $\omega$ if $X(\omega) \geq s$, that's because there are at most $rs$-trials that we can point to as a "certificate". More precisely, there are $\leq rs$ indices of $\omega$, so that for all other samples $\omega'$ that agree with $\omega$ on those $\leq rs$ indices, $X(\omega' \geq s)$ also.*

*Then, for all $t > \sqrt{E(X)}$, if $E(X)$ is sufficiently large and $\beta < 1/8c^2 r$, we have:*

$$\mathbb{P}(|X - \mathbb{E}(X)| > t) \leq 2e^{-\frac{\beta t^2}{\mathbb{E}(X)}}.$$

Note how the dependency on $n$ just vanished under the mild certifiability assumption, bringing us back to the quantitative strength of Chernoff, just magic!

What about random variables that are not determined by $n$ independent trials, but dependent trials? For example, this would be the case any time we want to analyse a random greedy process. We have some randomness at each step, but the randomness is also highly restricted by potentially bad choices we've made in the past. These kind of situations can turn into head-scratchers real fast, but the canonical way to deal with them is to reason about how much the information coming in at step $i$ would change our prediction of $X$ at any given step, and apply Azuma's inequality (the canonical martingale inequality):

**Lemma 4.6** (Azuma's inequality). *Let $X$ be a r.v. determined by **not necessarily independent** outcomes of $n$ trials, $T_1, \ldots, T_n$, such that for any $i \in [n]$, and any two possible sequences of outcomes*

$$t_1, \ldots, t_{i-1}, t_i$$

*and*

$$t_1, \ldots, t_{i-1}, t_i'$$

*we have*

$$\left| \mathbb{E}\left[ X \mid \bigwedge_{1 \leq j \leq i} T_j = t_j \right] - \mathbb{E}\left[ X \mid \bigwedge_{1 \leq j \leq i-1} T_j = t_j \text{ and } T_i = t_i' \right] \right| \leq c_i$$

*then*

$$\mathbb{P}(|X - \mathbb{E}(X)| > t) \leq 2e^{t^2/(2\sum_{i \in [n]} c_i^2)}$$

Azuma and McDiarmid are incomparable. Azuma lets us deal with dependencies across trials, and McDiarmid lets us get rid of the dependence on the number of trials, assuming independence and certifiability.

> **Exercise.** Derive McDiarmid's inequality from Azuma's inequality.

## 4.2 Defining and analysing the nibble

Recall that every bipartite $d$-regular graph has a perfect matching. Can we recover (an approximation of) this fact with the probabilistic method? It's not going to work if each vertex on one side chooses a neighbour uniformly at random. This doesn't even work for $K_{n,n}$ — with high probability, something like a third of vertices will not be seen by the candidate matching, i.e. there will be a lot of clashes (exercise!). Something that seems a lot more promising is the "random greedy algorithm". That is, at stage $i$, pick a random edge among all those edges that don't intersect with the previously chosen edges. This one works for trivial reasons in $K_{n,n}$, but does it work well for some arbitrary $d$-regular $2n$-vertex bipartite graph? (We may even imagine $d = \text{poly}(n)$ for simplicity.) In particular, can such an algorithm keep going, with high probability, for $(1 - o(1))n$ steps? The answer in this case is yes, and it can be shown by setting things up properly for an application of Azuma's inequality — but this it it fairly delicate to set things up properly and control the potential deviations in the conditional expectations *dynamically* along the whole trajectory of the random process.

The idea of nibble is extremely simple: let us do something in-between the stupid random algorithm of randomising all at once fully independently and the delicate random greedy algorithm: Let's randomise in small batches (bites) and slowly, but still just in $O(1)$ steps, eat away our graph with a partial matching until it's mostly gone. The idea is often associated to Rödl's work on constructing "approximate" designs, but goes back further than that to work of Ajtai, Komlós, Pintz, Spencer, and, you've guessed it, Szemerédi. It seems that Desmond Tutu has also discovered the basic idea independently, as the following quote suggests:

> **The nibble method**
> - "There's only one way to eat an elephant: one bite at a time."

Digesting the following definition is important for our framework.

> **Notation** Given real numbers $a, b, c$, we write $a = b \pm c$ to denote that $b - c \leq a \leq b + c$. We say a graph $G$ is $(1 \pm \gamma)d$-regular if $d_G(v) = (1 \pm \gamma)d$ for each $v \in V(G)$. We use standard notation for "hierarchies" of constants, writing $x \ll y$ to mean that there is a non-decreasing function $f : (0, 1] \to (0, 1]$ such that all the relevant subsequent statements hold for $x \leq f(y)$. Hierarchies with multiple constants are defined similarly, with the functions chosen from right to left, for example, for $x \ll y \ll z$.

With Tutu's idea, we can easily prove the following, which is a baby case of the nibble:

**Proposition 4.7.** *Suppose $d \geq n^{1/100}$, and suppose $1/n \ll \gamma \ll \varepsilon \leq 1$. Let $G$ be a $2n$-vertex graph that is $(1 \pm \gamma)d$-regular (all degrees are in this range). Then, $G$ has a matching of size $(1 - \varepsilon)n$.*

Recall the $\ll$ notation defined above: this is saying for every $\varepsilon$, if $\gamma$ is sufficiently small w.r.t. to $\varepsilon$, and $n$ is sufficiently large w.r.t. everything else, the assertion holds. We suppress the literal dependencies to highlight the hierarchy, which is what matters the most.

Proposition 4.7 can be derived from Hall's marriage theorem very easily, with linear depen-

dence between $\gamma$ and $\varepsilon$, and not much of an assumption about $n$ or $d$ in relation to $\gamma$ and $\varepsilon$. So why are we bothering with these trivialities? It's conceptually interesting (and has non-trivial applications) that a random algorithm can recover this, and as we will see, the proof generalises to higher dimensions, whereas Hall's marriage theorem does not.

We can derive the above from:

**Lemma 4.8.** *Suppose $d \geq n^{1/1000}$, and suppose $1/n \ll \gamma_1 \ll \gamma_2, \beta \leq 1$. Let $G$ be a $2n$-vertex graph that is $(1 \pm \gamma_1)d$-regular. Then, $G$ has a set of edges $M$ such that the following all hold.*

1. *$|M| = (1 \pm \gamma_2)\beta n$*
2. *$|V(M)| = (1 \pm \gamma_2)e^{-\beta}(2n)$*
3. *$G \setminus V(M)$ is a $(1 \pm \gamma_2)e^{-\beta}d$-regular graph.*

Note above $M$ is not necessarily a matching but close enough as it spans almost twice its size as a vertex set. $M$ will simply be a set of edges with each edge sampled with probability $\beta/d$, which is why the $e^{-\beta}$ terms pop up naturally.

> **Exercise.** Derive Proposition 4.7 from Lemma 4.8. How should we select $\beta$ in terms of the $\varepsilon$? Should we select the same $\beta$ each time? How should we set up the $\gamma_1, \gamma_2, \gamma_3 \ldots$. How many times should we plan on applying Lemma 4.8 as a function of $\varepsilon$? Is it an issue that the $M$ that the lemma gives isn't a matching?

**Proof of Lemma 4.8.** We will select $p = \beta/d$ and create $M$ by sampling each edge of $G$ at random, independently, with probability $p$.

Some basic calculations:

- $\mathbb{E}[|M|] = p \cdot e(G)$ (linearity of expectation) which is $= (1 \pm \gamma_1)pnd$. By independence, we can apply Chernoff to deduce $\mathbb{P}(|M| \neq (1 \pm 2\gamma_1)\beta n) \leq 2e^{-\beta n \gamma_1^2/3} = o_{n,d}(1)$.

- For any $v \in V(G)$, we have

$$\mathbb{P}(v \in V(M)) = 1 - (1-p)^{d_G(v)} \simeq 1 - e^{-pd_G(v)} = 1 - e^{-\beta(1 \pm 3\gamma_1)} = 1 - e^{-\beta}(1 \pm 5\gamma_1)$$

.

- For any subset $S \subseteq V(G)$, consider $I_S = \sum_{v \in S} I_{v \in V(M)} = |S \cap V(M)|$. If $S$ is restricted to come from just one of the sides of the bipartition, $I_S$ is a sum of Bernoulli random variables, as the random trials determining $I_{v \in V(M)}$ for distinct $v$ occurs on disjoint sets of edges which we sample independently! So we can apply Chernoff for sets of the form $S = N(v)$, which will be key to maintain regularity. So for any $v$ we have:

$$\mathbb{P}(|N(v) \setminus V(M)| \neq (1 \pm 10\gamma_1)e^{-\beta}) \leq e^{-\text{poly}(d)}$$

- By the same reasoning as above, we can deduce that

$$\mathbb{P}(|V(M)| \neq (1 \pm \gamma_2)e^{-\beta}(2n)) \leq e^{-\text{poly}(d)}$$

So with high positive probability, $M$ satisfies all of the above, and for each $v$ (check the union bound makes sense!). $\qquad\square$

> **Exercise.** What if the graph wasn't bipartite, but still $(1 \pm \gamma)d$-regular. Then Chernoff doesn't apply everywhere it should. How should we proceed? Does McDiarmid work? If we go for Talagrand, how can we think of the random variables so that they are "certifiable"?

### 4.3 The actual strong 99%-lemma that we need

For a 3-uniform, 3-partite hypergraph $H$, vertices $u, v$ and a subset $U \subseteq V(H)$, we define the **pair degree** of $(u, v)$ into $U$ as the number of vertices in $U$ which are in the neighbourhood of both $u$ and $v$, i.e. the number of vertices $z$ in $U$ such that there exists $v, w \in V(H)$ such that $\{u, z, v\}$ and $\{v, z, w\}$ are both edges of $H$. For a 2-uniform, bipartite graph $H$, vertices $u, v$, and a subset $U \subseteq V(H)$, we define the **pair degree** of $(u, v)$ into $U$ as the number of mutual neighbours of $u$ and $v$ in $U$.

We say that a $r$-partite $r$-uniform hypergraph $H$ (where $r$ will be either 2 or 3) is $(\gamma, p, n)$-**regular** if every part has $(1 \pm \gamma)n$ vertices and every vertex has degree $(1 \pm \gamma)pn$. We say that $H$ is $(\gamma, p, n)$-**typical** if, additionally, every pair of vertices $x, y$ in the same part of $H$ have pair degree $(1 \pm \gamma)p^2 n$ into every other part of $H$. We say that a hypergraph is **linear** if through every pair of vertices, there is at most one edge.

**Observation 4.9.** *Tripartite hypergraphs that are $(0, 1, n)$-typical and linear are exactly the hypergraphs that correspond to order $n$ Latin squares.*

One result we need is the following, which is proved by the above nibble idea if we keep track of the parameters.

**Lemma 4.10.** *Let $1/n \ll \gamma, \delta$ Every $(\gamma, \delta, n)$-regular linear tripartite hypergraph has a matching covering all but at most $n^{1-1/500} + 3\gamma n$ vertices.*

**Proof.** For an explicit proof, I recommend the Molloy–Reed book, or Luke Postle's lectures on youtube. $\qquad\square$

We also need an elementary property of typical bipartite graphs that follow from just a counting result. Typical graphs have the following well-known pseudorandomness condition, which dates back to work of Thomason [Tho89]. Note that for the rest of the section, we assume that bipartite graphs come with a partition of their vertex set as $(A, B)$ and similarly tripartite hypergraphs come with a partition $(A, B, C)$.

**Lemma 4.11.** *Let $G$ be a $(\gamma, \delta, n)$-typical bipartite graph. Then for every $A' \subseteq A, B' \subseteq B$, we have $e(A', B') = \delta|A'||B'| \pm 5\sqrt{(\delta + \gamma)n^3} + \gamma n^2$.*

Combined with the nibble, we can deduce the following, which will let us find a large matching even after an absorber is removed:

**Lemma 4.12.** *Let $H = H_G$ be a multiplication hypergraph. Let $p \geq n^{-1/650}$. Let $A', B', C'$ be $p$-random subsets of $A, B, C$ respectively, not necessarily independent. Set $R := A' \cup B' \cup C'$. With high probability the following holds. Let $q \leq 5p$. For any $X \subseteq V(H) \setminus R$ with $|X \cap A|, |X \cap B|, |X \cap C| = (1 \pm n^{-0.25})qn$, there is a matching in $R \cup X$ covering all but at most $n^{1-10^{-4}}$ vertices of $R \cup X$.*

# References

[BS21]    MATIJA BUCIC and BENNY SUDAKOV (2021). Tight ramsey bounds for multiple copies of a graph. *arXiv preprint arXiv:2108.11946* . ↑3

[Kee14]   PETER KEEVASH (2014). The existence of designs. *arXiv preprint arXiv:1401.3665* . ↑1, 4

[LSM24]  RICHARD LANG and NICOLÁS SANHUEZA-MATAMALA (2024). A hypergraph bandwidth theorem. *arXiv preprint arXiv:2412.14891* . ↑8, 9

[LW17]    ANITA LIEBENAU and NICK WORMALD (2017). Asymptotic enumeration of graphs by degree sequence, and the degree sequence of a random graph. *arXiv preprint arXiv:1702.08373* . ↑10

[Mon]     R. MONTGOMERY. A proof of the Ryser-Brualdi-Stein conjecture for large even $n$ ArXiv preprint 2310.19779. ↑1, 4, 5

[Rei89]   K. B. REID (1989). Three problems on tournaments. *Graph Theory and its Applications: East and West, Annals of the New York Academy of Sciences*, vol. 576, 466–473. ↑1

[RRS06]   VOJTĚCH RÖDL, ANDRZEJ RUCIŃSKI, and ENDRE SZEMERÉDI (2006). A Dirac-type theorem for 3-Uniform hypergraphs. *Combinatorics, Probability and Computing* **15**(1-2), 229–251. ↑1, 8, 12

[SL22]    MICHAEL SIMKIN and ZUR LURIA (2022). A lower bound for the n-queens problem. *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2185–2197 (SIAM). ↑1

[Tho89]   ANDREW THOMASON (1989). Dense expanders and pseudo-random bipartite graphs. *Annals of Discrete Mathematics*, vol. 43, 381–386. ↑18