Probabilistic graph inspections through forests

Koperberg, V.T.

Citation

Koperberg, V. T. (2026, June 25). *Probabilistic graph inspections through forests*. Retrieved from https://hdl.handle.net/1887/4307047

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: https://hdl.handle.net/1887/4307047

Note: To cite this publication please use the final published version (if applicable).

PART II

COUPLINGS AND MATCHINGS

CHAPTER 5

Couplings and Matchings

This chapter is based on the following paper: T. Koperberg. “Couplings and Matchings: Combinatorial notes on Strassen’s theorem”. In: *Statistics & Probability Letters* 209 (2024), p. 110089.

Abstract

Some mathematical theorems represent ideas that are discovered again and again in different forms. One such theorem is Hall’s marriage theorem. This theorem is equivalent to several other theorems in combinatorics and optimization theory, in the sense that these results can easily be derived from each other. Remarkably, this equivalence extends to Strassen’s theorem, a celebrated result on couplings of probability measures.

In this paper the equivalence between Strassen’s theorem and Hall’s theorem is investigated from a combinatorial perspective. A novel combinatorial lemma will be introduced that can be used to deduce both Hall’s theorem and a finite version of Strassen’s theorem, providing a simple proof of their equivalence.

5.1 Introduction

In the original paper from 1935 Hall already mentions a similarity between his *marriage theorem* and a result by König from 1916. Since then numerous other results have been found that are 'equivalent' to Hall's theorem. This equivalence is an informal concept and simply means that two results can be derived from each other via simple proofs. This class of equivalent theorems includes among others *Menger's theorem* (1927), *König's minimax theorem* (1931), the *Birkhoff-von Neumann theorem* (1946), *Dilworth's theorem* (1950) and the *max-flow min-cut theorem* by Ford and Fulkerson (1956). An extensive discussion on these equivalences can be found in [73].

Another result that belongs to this class of equivalent statements is *Strassen's theorem* (1965). As this theorem is a result from probability theory, the original proof made use of analytical tools rather than the combinatorial methods used in the proofs of the above mentioned theorems. Therefore, it is remarkable that this result is, in fact, equivalent to these combinatorial statements. This relation between Strassen's theorem and Hall's theorem is already known in the literature, as Dudley used Hall's marriage theorem to prove an extension of Strassen's original result [23].

In this paper we will look at Strassen's theorem from a combinatorial perspective. Our discussion will be restricted to a finite variant of Strassen's theorem, which is stated in theorem 5.1. For the general version of the theorem the reader is referred to [57].

The goal of this paper is twofold: firstly to give a combinatorial proof of the finite version of Strassen's theorem directly from first principles, and secondly to give a simple proof of the equivalence between Strassen's theorem and Hall's theorem, which will be done using an adaptation of Dudley's proof. For both of these objectives we will make use of a novel lemma, that will be introduced in section 5.2.1, and which will be referred to as the *subforest lemma*. As will be discussed in remark 5.4, this lemma could be derived from a more abstract result within the theory of optimal transport. In section 5.1.1 we will introduce the two main theorems. The content of the subsequent sections is outlined in fig. 5.1.

5.1.1 The two main theorems

We will start by introducing the two theorems that are the main topic of this paper.

If $\mathbb{P}$ and $\mathbb{P}'$ are probability measures on two finite sets A and B, respectively, then a *coupling* of $\mathbb{P}$ and $\mathbb{P}'$ is any probability measure $\hat{\mathbb{P}}$ on the product set $A \times B$ for which its marginals correspond to $\mathbb{P}$ and $\mathbb{P}'$. That is, for all $U \subseteq A$ and $S \subseteq B$ it holds that $\mathbb{P}(U) = \hat{\mathbb{P}}(U \times B)$ and $\mathbb{P}'(S) = \hat{\mathbb{P}}(A \times S)$.

Theorem 5.1 (Strassen's theorem for finite sets). *Let A and B be finite sets and $R \subseteq A \times B$ a relation between them. Let $\mathbb{P}$ and $\mathbb{P}'$ be probability measures on A and B, respectively. Then there exists a coupling $\hat{\mathbb{P}}$ of $\mathbb{P}$ and $\mathbb{P}'$ with $\hat{\mathbb{P}}(R) = 1$ if and only if*

$$\mathbb{P}(U) \leq \mathbb{P}'(N_R(U)), \qquad \text{for all } U \subseteq A, \tag{5.1}$$

where $N_R(U) = \{y \in B \colon\ \exists x \in U \text{ s.t. } (x, y) \in R\}$.

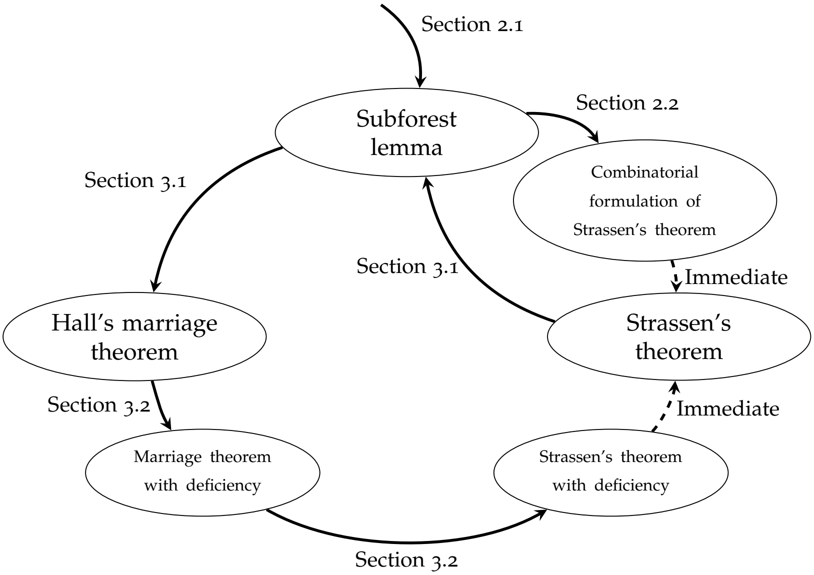

Figure 5.1: A graphical outline of this paper, where the arrows represent the different proofs.

We will refer to (5.1) as the *coupling condition.* In [26] it is shown how the general version of Strassen's theorem can be derived from this finite version.

In this paper we will use the graph theoretic formulation of the marriage theorem. All graphs in this paper are assumed to be simple finite undirected graphs. A *bipartite graph* is a graph of which the vertices can be partitioned into two sets $\{A, B\}$ such that all edges have one endpoint in A and the other endpoint in B. This partition $\{A, B\}$ will be called the *bipartition* of the graph. A *matching* of a graph is a subset M of its edges such that all vertices are incident to at most one edge in M. If all vertices are incident to an edge in M, then M is called a *perfect matching.*

Theorem 5.2 (Hall's marriage theorem). *Let G be a bipartite graph with bipartition $\{A, B\}$ such that $|A| = |B|$. Then G contains a perfect matching if and only if it holds that*

$$|U| \leq |N_G(U)|, \qquad \text{for all } U \subseteq A. \tag{5.2}$$

Here $N_G(U)$ denotes the set of vertices that are neighbors of vertices in U. If the underlying graph is clear, then the subscript will be dropped. We will refer to (5.2) as the *marriage condition.*

5.2 Independent proof of Strassen's theorem

5.2.1 The subforest lemma

A graph that does not contain any cycles is called a *forest.* A *weighted graph* is a graph that is equipped with a vertex weight function $w : V \to [0, \infty)$. For such weight functions we write $w(U) = \sum_{x \in U} w(x)$ for $U \subseteq V$. Unless otherwise specified, a subgraph of a weighted graph is equipped with the restriction of the weight function to the vertices of the subgraph. So, in particular a spanning subgraph has the same weight function as the underlying full graph. For brevity we will call a spanning subgraph a *subforest* when it is a forest.

Lemma 5.3 (Subforest lemma). *Let $G = (V, E, w)$ be a weighted bipartite graph with bipartition $\{A, B\}$ such that $w(A) = w(B)$. If it holds that*

$$w(U) \leq w(N_G(U)), \quad \text{for all } U \subseteq A, \tag{5.3}$$

then G contains a subforest that satisfies (5.3).

We will refer to (5.3) as the *subforest condition.* Both the marriage condition (5.2) and the coupling condition (5.1) are special cases of this subforest condition. For the marriage condition all vertices have unit weight, while for the coupling condition the weight function is normalized so that $w(A) = w(B) = 1$. Note that these three conditions seem to break the symmetry between sets A and B that is present in the setting of the theorems. This is in fact not the case, as it can be easily verified that (5.3) implies that $w(U) \leq w(N_G(U))$ for all $U \subseteq B$.

Here we give an independent proof of lemma 5.3 directly from first principles. The proof uses the same strategy used in the inductive proof of the marriage theorem by Halmos and Vaughan [34], in which the induction hypothesis acts as a marriage broker. That is, we distinguish between the case where the graph contains a 'critical set' of vertices and the case where no such set exists.

Proof of lemma 5.3. We will apply induction on $|V|$. If $|V| = 2$, then G is itself a forest, so there is nothing to prove. Now assume that $|V| > 2$ and that the statement holds when $|V|$ is smaller. Let $\mathcal{S} = \{U \subseteq A\colon\ 0 < |U| < |A|\} \cup \{U \subseteq B\colon\ 0 < |U| < |B|\}$ denote the collection of non-empty strict subsets of either A or B. We will distinguish two cases.

In the first case we assume that there exists a $U \in \mathcal{S}$ with $w(U) = w(N_G(U))$. Without loss of generality we assume that $U \subseteq A$. Let $\{V_1, V_2\}$ be the partition of V given by $V_1 = U \cup N_G(U)$ and $V_2 = (A \setminus U) \cup (B \setminus N_G(U))$. Then both induced subgraphs $G[V_1]$ and $G[V_2]$ satisfy the subforest condition (5.3). Thus by the induction hypothesis there exist subforests F_1 and F_2 of $G[V_1]$ and $G[V_2]$, respectively, that both satisfy the subforest condition (5.3). The graph $F = (V, E(F_1) \cup E(F_2))$, that contains all edges of F_1 and F_2, is a subforest of G, that satisfies the subforest condition with respect to w. Also note that F contains at least two connected components, since none of the vertices in V_1 is connected to any of the vertices in V_2.

For the second case we assume that $w(U) < w(N_G(U))$ for all $U \in \mathcal{S}$. Let ε denote the minimal weight of any vertex of G, i.e. $\varepsilon = \min_{v \in V} w(v)$. Let $x \in V$ be any vertex with $w(x) = \varepsilon$. Without loss of generality we can assume that $x \in A$. Let $y \in N_G(x)$ be any neighbor of x. Let $\mathcal{U} = \{U \subseteq A\colon\ x \notin U \text{ and } U \cap N_G(y) \neq \varnothing\}$ and take

$$\delta = \min_{U \in \mathcal{U}} w(N_G(U)) - w(U).$$

Since $w(N_G(y)) > w(y) \geq w(x)$, vertex y has at least two neighbors. It follows that $A - x \in \mathcal{U}$ and thus that

$$\delta \leq w(N_G(A-x)) - w(A-x) = w(B) - w(A-x) = \varepsilon.$$

Let $D \in \mathcal{U}$ be such that $w(N_G(D)) - w(D) = \delta$. We add a new element $\tilde{x}$ to A to obtain $\tilde{A} = A + \tilde{x}$. Let $\tilde{V} = V + \tilde{x}$, $\tilde{E} = E(G) + \{\tilde{x}, y\}$ and $\tilde{G} = (\tilde{V}, \tilde{E})$. Define the weight function $\tilde{w}$ on $V + \tilde{x}$ by $\tilde{w} = w + \delta \mathbf{1}_{\{\tilde{x}\}} - \delta \mathbf{1}_{\{x\}}$. The weighted graph $(\tilde{G}, \tilde{w})$ now satisfies the subforest condition, since for all $U \subseteq A - x$ with $y \in N_G(U)$ it holds that $U \in \mathcal{U}$, so that

$$\tilde{w}(U + \tilde{x}) = w(U) + \delta \leq w(N_G(U)) = \tilde{w}(N_{\tilde{G}}(U + \tilde{x})),$$

while if $y \notin N_G(U)$, then

$$\tilde{w}(U + \tilde{x}) \leq w(N_G(U)) + \delta = \tilde{w}(N_{\tilde{G}}(U + \tilde{x})) - \tilde{w}(y) + \delta \leq \tilde{w}(N_{\tilde{G}}(U + \tilde{x})).$$

For subsets $U \subseteq A$ that include x or that do not include $\tilde{x}$ condition (5.3) is also easily verified. We now have that $\tilde{w}(D + \tilde{x}) = \tilde{w}(N_{\tilde{G}}(D + \tilde{x}))$, so that $(\tilde{G}, \tilde{w})$ satisfies the assumptions of the first case.

If $D \neq A - x$, then $|(D + \tilde{x}) \cup N_{\tilde{G}}(D + \tilde{x})| < |V|$. It follows from the induction hypothesis, in the same manner as in the previous case, that there exists a subforest $\tilde{F}$ of $\tilde{G}$ with x and y in two distinct components such that $(\tilde{F}, \tilde{w})$ satisfies (5.3). The graph $F = (V, E(\tilde{F}) - \{\tilde{x}, y\} + \{x, y\})$ is a spanning subgraph of G. Since x and y are contained in distinct components of $\tilde{F}$, we also have that F is a forest. It is also clear that (F, w) satisfies the subforest condition.

If instead we have that $D = A - x$, then $\varepsilon = \delta$. Define the weight function w' on $V - x$ by $w' = w - \delta \mathbf{1}_{\{y\}}$. Then the weighted graph $(G[V-x], w')$ satisfies the subforest condition, since for all $U \subseteq A - x$ with $y \in N_G(U)$ it holds that $U \in \mathcal{U}$, so that

$$w'(U) = w(U) \leq w(N_G(U)) - \delta = w'(N_{G[V-x]}(U)),$$

while if $y \notin N_G(U)$, then

$$w'(U) = w(U) \leq w(N_G(U)) = w'(N_{G[V-x]}(U)).$$

Hence, by the induction hypothesis, there exists a spanning subforest F' of $G[V-x]$ satisfying the subforest condition. Let $F = (V, E(F') + \{x, y\})$. Then F is a subforest of G satisfying the subforest condition.

In both cases we have shown the existence of a spanning subforest that satisfies the subforest condition, thus completing the proof. $\square$

Remark 5.4. The problem of finding a coupling that satisfies the coupling condition (5.1) can also be phrased as an optimal transport problem. A solution to such a transportation problem corresponds to a bipartite graph with weights assigned to the edges. Klee and Witzgall [48] showed that the polytope of feasible solutions has at its vertices exactly those solutions whose accompanying bipartite graph corresponds to a forest. Hence, the subforest lemma can also be derived from this result of Klee and Witzgall.

5.2.2 Deriving Strassen's theorem from the subforest lemma

For the independent proof of Strassen's theorem for finite sets, we show how it can easily be derived from the subforest lemma. It is natural to translate the setting of theorem 5.1 to a weighted bipartite graph $G = (V, E, w)$ defined by

$$V = A \cup B, \quad E = \{\{x, y\} \colon \ (x, y) \in R\}, \text{ and}$$
$$w(x) = \begin{cases} \mathbb{P}(x) & \text{if } x \in A \\ \mathbb{P}'(x) & \text{if } x \in B \end{cases} \tag{5.4}$$

(Here we assume w.l.o.g. that $A \cap B = \varnothing$.) The coupling condition then translates to $w(U) \leq w(N_G(U))$ for all $U \subseteq A$, while the sought coupling becomes an edge weight function $\hat{w} : E \to [0, \infty)$ that satisfies $w(x) = \sum_{e \sim x} \hat{w}(e)$ for all $x \in A$, where the sum is taken over all edges incident to x. Note that the vertex weight function w obtained in this manner will always satisfy $w(A) = w(B) = 1$, due to the fact that probability measures have total mass 1. This translation, which is further illustrated in fig. 5.2, gives us the following equivalent formulation of theorem 5.1, which resembles a weighted version of Hall's marriage theorem. Theorem 5.1 follows directly from proposition 5.5 by normalizing the vertex and edge weights, so that these form probability measures.

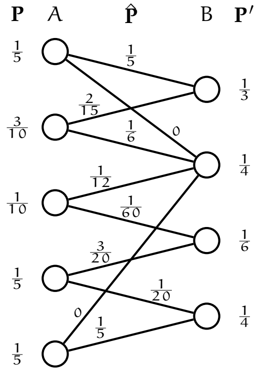

Figure 5.2: A graphical representation of the setting of theorem 5.1. The vertices represent the elements of sets A and B, and the edges represent the elements of R. The number next to a vertex denotes the mass that either $\mathbb{P}$ or $\mathbb{P}'$ assigns to that vertex. The numbers above the edges are the masses assigned by the coupling $\hat{\mathbb{P}}$, which is supported on a subset of R. By interpreting the masses of $\mathbb{P}$ and $\mathbb{P}'$ as vertex weights and the masses of $\hat{\mathbb{P}}$ as edge weights, we find ourselves in the setting of proposition 5.5.

Proposition 5.5 (Combinatorial formulation of Strassen's theorem). *Let $G = (V, E, w)$ be a weighted bipartite graph with bipartition $\{A, B\}$ such that $w(A) = w(B)$. Then the following are equivalent:*

- *(i) for all $U \subseteq A$ it holds that $w(U) \leq w(N(U))$;*
- *(ii) there exists an edge weight function $\hat{w} : E \to [0, \infty)$ such that for all $x \in V$ it holds that $w(x) = \sum_{e \sim x} \hat{w}(e)$, where the sum is taken over all edges incident to x.*

Proof of proposition 5.5 using lemma 5.3. The implication from ((ii)) to ((i)) is easily shown. The set of edges with one endpoint in U is a subset of the set of edges with one endpoint in $N(U)$. Hence, if $\hat{w}$ satisfies ((ii)), then

$$w(U) = \sum_{x\in U}\sum_{e\sim x}\hat{w}(e) \leq \sum_{y\in N(U)}\sum_{e\sim y}\hat{w}(e) = w(N(U)).$$

The reverse implication will be proven by induction on $|V|$. Since w satisfies ((i)), by the subforest lemma there exists a subforest F of G satisfying ((i)). Since F is a forest, there exists a vertex x in F with degree 1. Without loss of generality we can assume that $x \in A$. Let $y \in B$ be the unique neighbor of x in F.

Note that it follows from ((i)) that $w(x) \leq w(y)$. Set $\varepsilon = w(x)$. Consider the induced subgraph $F[V-x]$ obtained by removing vertex x from F and equip it with the vertex weight function $\tilde{w} : V - x \to [0,\infty)$ given by $\tilde{w}(v) = w(v) - \varepsilon\mathbf{1}_{\{v=y\}}$. The weighted graph $(F[V-x], \tilde{w})$ satisfies ((i)), since for all $U \subseteq A - x$ with $y \in N_F(U)$ it holds that

$$\tilde{w}(U) = w(U) = w(U+x) - \varepsilon \leq w(N_F(U+x)) - \varepsilon = w(N_F(U)) - \varepsilon = \tilde{w}(N_{\tilde{F}}(U)),$$

while for all $U \subseteq A - x$ with $y \notin N_F(U)$ it holds that

$$\tilde{w}(U) = w(U) \leq w(N_F(U)) = \tilde{w}(N_{\tilde{F}}(U)).$$

Hence, by the induction hypothesis there exists an edge weight function $\hat{w}$ on $F[V-x]$ satisfying ((ii)). Now define an edge weight function on the edges of G by

$$\breve{w}(e) = \begin{cases} \hat{w}(e) & \text{if } e \in E(F[V-x]) \\ \varepsilon & \text{if } e = \{x,y\} \\ 0 & \text{otherwise.} \end{cases}$$

Then $\breve{w}$ is the sought edge weight function satisfying ((ii)). $\square$

This independent proof of Strassen's theorem for finite sets is constructive and can in principle be used to find the required coupling. However, far more efficient methods for finding such a coupling exists. In [58, Corollary 2.1.5] it is mentioned that proposition 5.5 can be derived from the max-flow min-cut theorem, see also [60, Theorem 10.4]. The method is similar to the derivation of the marriage theorem from the max-flow min cut theorem, that is given in [27]. This derivation is not only elegant, it also shows that any method for finding maximal network flows can also be used to find such a coupling.

5.3 Equivalence of Hall's theorem and Strassen's theorem

In the second part of this paper we prove the equivalence of Strassen's theorem for finite sets and Hall's marriage theorem.

5.3.1 Deriving Hall's theorem from Strassen's theorem

The derivation of Hall's theorem from Strassen's theorem will go via the subforest lemma.

Proof of lemma 5.3 using Strassen's theorem. The statement will be proven by induction on $|E|$. If $|E| = 1$, then G is itself a forest. Now assume that $|E| \geq 2$ and that the statement holds when $|E|$ is smaller.

Define the probability measures $\mathbb{P}$ and $\mathbb{P}'$ on A and B, respectively, by setting $\mathbb{P}(x) = \frac{w(x)}{w(A)}$ and $\mathbb{P}(y) = \frac{w(y)}{w(B)}$ for $x \in A$ and y in B. Since G satisfies the subforest condition, these two probability measures then satisfy the coupling condition with respect to the relation E. Hence, by Strassen's theorem there exists a coupling $\hat{\mathbb{P}}$ of $\mathbb{P}$ and $\mathbb{P}'$ that is supported on E.

If G is not a forest, then there is a subset of edges $C \subseteq E$ that constitute a cycle. Now take $\varepsilon = \min\{\hat{\mathbb{P}}(e)\colon\ e \in C\}$ and let $e^* \in C$ be such that $\hat{\mathbb{P}}(e^*) = \varepsilon$. Since G is a bipartite graph, the cycle C contains an even number of edges. Hence, we can partition C into two sets $\{I_C, J_C\}$ such that edges in I_C are only incident to edges in J_C and vice-versa. Without loss of generality we can assume that $e^* \in I_C$.

Now define a new probability measure $\tilde{\mathbb{P}}$ on E by

$$\tilde{\mathbb{P}}(e) = \begin{cases} \hat{\mathbb{P}}(e) - \varepsilon & \text{if } e \in I_C \\ \hat{\mathbb{P}}(e) + \varepsilon & \text{if } e \in J_C \\ \hat{\mathbb{P}}(e) & \text{otherwise.} \end{cases}$$

Since each vertex in G is incident to the same number of edges in I_C as to edges in J_C, we have that $\tilde{\mathbb{P}}$ is also a coupling of $\mathbb{P}$ and $\mathbb{P}'$. Moreover, the coupling $\tilde{\mathbb{P}}$ is supported on $E \setminus \{e^*\}$, since by construction it holds that $\tilde{\mathbb{P}}(e^*) = 0$. Thus, by applying Strassen's theorem in the reverse direction, we find that the relation $E \setminus \{e^*\}$ satisfies the coupling condition with respect to $\mathbb{P}$ and $\mathbb{P}'$. It follows that the weighted graph $G - e^*$ satisfies the subforest condition. By the induction hypothesis $G - e^*$ contains a subforest F satisfying that condition. Clearly, F is also a subforest of G, which finishes the proof. ☐

Proof of theorem 5.2 using the subforest lemma. We will only prove the sufficiency of the marriage condition, which will be done by induction on $|E|$. So, we assume that G satisfies the marriage condition.

Clearly the statement holds if $|E| = 1$. Now assume that $|E| \geq 2$ and that the statement holds if $|E|$ is smaller. Note that the marriage condition is a special case of the subforest condition where each vertex has unit weight. Hence, by lemma 5.3 there exists a subforest F of G that satisfies the marriage condition. Since F is a forest, there exists a vertex x in F with degree 1. Let y be the unique neighbor of x in F. Then the induced subgraph $G[V \setminus \{x, y\}]$ still satisfies the marriage condition. Thus by the induction hypothesis $G[V \setminus \{x, y\}]$ has perfect matching M. Taking $M \cup \{x, y\}$ gives a perfect matching of G. ☐

5.3.2 Deriving Strassen's theorem from the marriage theorem

To finish our reciprocal derivations we still have to prove Strassen's theorem from the marriage theorem. This will be done using two well-known generalizations of both theorems, propositions 5.6 and 5.7 below, that allow for some small deficiencies in the conditions.

The used generalization of Hall's marriage theorem is due to Ore [66] and can be found in e.g. [58, Thm. 1.3.1]. It can be easily derived from the marriage theorem itself, which led Mirsky to call the marriage theorem a *self-refining result* [65]. For completeness we also give this derivation.

Proposition 5.6 (Hall's theorem with deficiency). *Let G be a bipartite graph with bipartition $\{A, B\}$ with $|A| = |B| = n$. Then G contains a matching M with $|M| \geq n - k$ if and only if it holds that*

$$|U| \leq |N_G(U)| + k, \qquad \text{for all } U \subseteq A. \tag{5.5}$$

Proof of proposition 5.6 using Hall's marriage theorem. We only prove the sufficiency of (5.5).

Construct the bipartite graph $\tilde{G}$ by adding k new vertices $a_1, \ldots, a_k$ to A and k new vertices $b_1, \ldots, b_k$ to B. Set $\tilde{A} = A \cup \{a_1, \ldots, a_k\}$ and $\tilde{B} = B \cup \{b_1, \ldots, b_k\}$. Also add edges between all a_i and all vertices in $\tilde{B}$ and all b_i and all vertices in $\tilde{A}$. That is, $\tilde{G} = (\tilde{A} \cup \tilde{B}, \tilde{E})$ with $\tilde{E} = E \cup \{\{a_i, v\}\colon\ 1 \leq i \leq k,\ v \in \tilde{B}\} \cup \{\{b_i, v\}\colon\ 1 \leq i \leq k,\ v \in \tilde{A}\}$.

Since G satisfies (5.5) and all vertices have k more neighbors in $\tilde{G}$ than in G, we have that $\tilde{G}$ satisfies the marriage condition (5.2). Hence, by theorem 5.2 $\tilde{G}$ contains a perfect matching $\tilde{M}$. Note that $|\tilde{M}| = n + k$. Hence, $M := \tilde{M} \cap E$ is a matching of G with $|M| \geq n - k$, since at most $2k$ edges in $\tilde{M}$ are incident to any of the $2k$ vertices that are not in G. ☐

As with proposition 5.6 the generalized version of Strassen's theorem also follows from its original. However, for our purposes we will derive it from proposition 5.6 instead. The proof is an adaptation of the proof by Dudley given in [24, Theorem 11.6.3] adjusted to our finite setting.

Proposition 5.7 (Strassen's theorem with deficiency). *Let A and B be finite sets and $R \subseteq A \times B$ a relation between them. Let $\mathbb{P}$ and $\mathbb{P}'$ be probability measures on A and B, respectively. Let $\varepsilon \geq 0$ be given. Then there exists a coupling $\hat{\mathbb{P}}$ of $\mathbb{P}$ and $\mathbb{P}'$ with $\hat{\mathbb{P}}(R) \geq 1 - \varepsilon$ if and only if*

$$\mathbb{P}(U) \leq \mathbb{P}'(N_R(U)) + \varepsilon, \qquad \text{for all } U \subseteq A. \tag{5.6}$$

Proof of proposition 5.7 using proposition 5.6. For the necessity of (5.6) we note that if $\hat{\mathbb{P}}$ is a coupling of $\mathbb{P}$ and $\mathbb{P}'$ with $\hat{\mathbb{P}}(R) \geq 1 - \varepsilon$, then it holds for all $U \subseteq A$ that

$$\mathbb{P}(U) = \hat{\mathbb{P}}(U \times B) \leq \hat{\mathbb{P}}(U \times N_R(U)) + \varepsilon \leq \hat{\mathbb{P}}(A \times N_R(U)) + \varepsilon = \mathbb{P}'(N_R(U)) + \varepsilon.$$

It remains to prove its sufficiency. This will be done in two steps. In the first step we assume that $\mathbb{P}$ and $\mathbb{P}'$ are both rational valued and that $\varepsilon \in \mathbb{Q}$, and in the second step we derive the result for arbitrary $\mathbb{P}$, $\mathbb{P}'$ and ε.

$\boxed{\text{Step (1)}}$ Assume that $\mathbb{P}$ and $\mathbb{P}'$ are rational valued and also take ε rational. Define $G = (V, E, w)$ as in (5.4). Then we have that $w(x) \in \mathbb{Q}$ for all $x \in V$. Since V is finite, there exists a large enough $N \in \mathbb{N}$ such that the product $Nw(x)$ is an integer for all $x \in V$ and such that $k := \varepsilon N$ is an integer as well.

Let $\tilde{V} = \bigcup_{x\in A} \bigcup_{i=1}^{Nw(x)} \{x_i\}$ be the set consisting of $Nw(x)$ copies of each element $x \in V$. Now consider the bipartite graph $\tilde{G} = (\tilde{V}, \tilde{E})$, where the edge set is given by $\tilde{E} = \{\{x_i, y_j\}\colon\ \{x, y\} \in E\}$. That is, two vertices x_i and y_j in $\tilde{G}$ are connected by an edge if and only if their originals x and y are adjacent in G. Denote the bipartition of $\tilde{G}$ by $\{\tilde{A}, \tilde{B}\}$.

Since $\mathbb{P}$ and $\mathbb{P}'$ satisfy (5.6), we then have for all $\tilde{U} \subseteq \tilde{A}$ that

$$|\tilde{U}| \le |N_{\tilde{G}}(\tilde{U})| + k.$$

By proposition 5.6 there exists a matching $\tilde{M}$ of $\tilde{G}$ with $|\tilde{M}| = N - k$. Let $a_1, \dots, a_k$ and $b_1, \dots, b_k$ denote the k vertices in $\tilde{A}$ and $\tilde{B}$, respectively, that are unmatched by $\tilde{M}$. Now consider the set of edges $M^+ := \tilde{M} \cup \{\{a_i, b_i\}\colon\ i \in [k]\}$, which is obtained from $\tilde{M}$ by adding k arbitrary edges, not necessarily belonging to $\tilde{E}$, between the unmatched vertices.

For each pair $(x, y) \in A \times B$ let

$$\hat{w}(x, y) = \left| \bigcup_{i=1}^{Nw(x)} \bigcup_{j=1}^{Nw(y)} \{x_i, y_j\} \cap M^+ \right|$$

denote the number of edges between copies of x and copies of y that occur in M^+. Since M^+ is a perfect matching of the complete bipartite graph on $\tilde{A} \cup \tilde{B}$, we find that $\sum_{y\in B} \hat{w}(x, y) = Nw(x)$ for all $x \in A$ and similarly that $\sum_{x\in A} \hat{w}(x, y) = Nw(y)$ for all $y \in B$. So, the probability measure $\hat{\mathbb{P}}$ on $A \times B$ defined by $\hat{\mathbb{P}}(x, y) = \frac{\hat{w}(x,y)}{N}$ is a coupling of $\mathbb{P}$ and $\mathbb{P}'$. Since only k of the edges of M^+ do not belong to $\tilde{M}$ we also find that $\hat{\mathbb{P}}(R) = 1 - \varepsilon$, so $\hat{\mathbb{P}}$ is the sought coupling.

$\boxed{\text{Step (2)}}$ Let $\mathbb{P}$, $\mathbb{P}'$ and ε be arbitrary. Let $(\varepsilon_i)_{i\in\mathbb{N}}$ be a rational sequence converging to ε from above. Since $\mathbb{P}$ and $\mathbb{P}'$ satisfy (5.6), we can find two sequences $(\mathbb{P}_i)_{i\in\mathbb{N}}$ and $(\mathbb{P}'_i)_{i\in\mathbb{N}}$ of rational valued probability measures that converge to $\mathbb{P}$ and $\mathbb{P}'$, respectively, such that for every $i \in \mathbb{N}$ it holds that

$$\mathbb{P}_i(U) \le \mathbb{P}'_i(N_R(U)) + \varepsilon_i, \qquad \text{for all } U \subseteq A.$$

By the first step of the proof, for each i there exists a coupling $\hat{\mathbb{P}}_i$ of $\mathbb{P}_i$ and $\mathbb{P}'_i$ with $\hat{\mathbb{P}}_i(R) \ge 1 - \varepsilon_i$. We can interpret $(\hat{\mathbb{P}}_i)_{i\in\mathbb{N}}$ as a sequence in the compact metric space $[0, 1]^{|E|}$. Thus it contains a converging subsequence $(\hat{\mathbb{P}}_{i_j})_{j\in\mathbb{N}}$ with limit $\hat{\mathbb{P}}$. It follows that $\hat{\mathbb{P}}(R) \ge 1 - \varepsilon$.

It remains to be shown that $\hat{\mathbb{P}}$ is a coupling of $\mathbb{P}$ and $\mathbb{P}'$. Let $\delta > 0$ be given. Then for all $x \in A$ there exists a $k \in \mathbb{N}$ such that for all $j \ge k$ it holds that both $|\mathbb{P}_{i_j}(x) - \mathbb{P}(x)| < \delta$ and

$$|\hat{\mathbb{P}}_{i_j}(\{x\} \times B) - \hat{\mathbb{P}}(\{x\} \times B)| < \delta.$$

It follows that

$$\begin{aligned} |\mathbb{P}(x) - \hat{\mathbb{P}}(\{x\} \times B)| &< |\mathbb{P}(x) - \hat{\mathbb{P}}_{i_k}(\{x\} \times B)| + \delta \\ &= |\mathbb{P}(x) - \mathbb{P}_{i_k}(x)| + \delta \\ &< 2\delta. \end{aligned}$$

Similarly, we find that $|\mathbb{P}'(x) - \hat{\mathbb{P}}(A \times \{x\})| < 2\delta$ for all $x \in B$. As this holds for all $\delta > 0$, it follows that $\hat{\mathbb{P}}$ is a coupling of $\mathbb{P}$ and $\mathbb{P}'$. □