Chapter 3

Matchings and related structures

A problem that has played a historically important role in the development both of complexity theory and algorithm design is that of evaluating the permanent function. The *permanent* of an $n \times n$ integer matrix $A = (a_{ij} : 0 \le i, j \le n-1)$ is defined by

$$\operatorname{per} A = \sum_{\pi} \prod_{i=0}^{n-1} a_{i,\pi(i)} ,$$

where the sum is over all permutations π of $[n] = \{0, \ldots, n-1\}$. Evaluating the permanent of a 0,1-matrix is complete for the class $\#\mathsf{P}$; thus, we cannot expect to find an algorithm that solves the problem exactly in polynomial time. Interest has therefore centred on finding computationally feasible approximation algorithms. In contrast, as is well known, the superficially related *determinant* of an $n \times n$ matrix can be evaluated in $\mathrm{O}(n^3)$ arithmetic operations using Gaussian elimination.

A *matching* in a graph $G = (V, E)$ is any subset $A \subseteq E$ of edges that are pairwise vertex disjoint. A matching is said to be *perfect* if it covers every vertex; clearly a perfect matching, which can exist only if $|V|$ is even, has size $|V|/2$. Specialised to the case when A is a 0,1-matrix, $\operatorname{per} A$ is equal to the number of perfect matchings in the bipartite graph $G = (V_1, V_2, E)$, where $V_1 = V_2 = [n]$, and $(i, j) \in E$ iff $a_{ij} = 1$.

In the light of the above connection, a promising approach to computing an approximation of the permanent of A, at least when A is a 0,1-matrix, is through sampling perfect matchings in the related bipartite graph G. In fact, it seems that neither the bipartiteness of G, nor the perfection of the matching is of crucial importance, so we shall immediately generalise the situation to that of sampling a (weighted) matching in a general graph. In the next section we attack that sampling problem through Markov chain simulation; then in subsequent sections we shall apply the methods we develop there to related problems, including the approximation of the permanent.

3.1 Weighted matchings (the monomer-dimer model)

Let G be a graph, not necessarily bipartite, with an even number $2n = |V|$ of vertices. The assumption that the number of vertices in G is even is inessential and is made for notational convenience. To each matching M, a *weight* $w(M) = \lambda^{|M|}$ is assigned, where λ is a positive real parameter. The generating (or partition) function of matchings in G is

$$Z(\lambda) \equiv Z_G(\lambda) = \sum_M w(M) = \sum_{k=0}^{n} m_k \lambda^k, \tag{3.1}$$

where $m_k \equiv m_k(G)$ is the number of k-matchings in G. In statistical physics, a matching is termed a "monomer-dimer configuration": the edges in M are the "dimers" and the unmatched (uncovered) vertices are "monomers". Thus $m_k(G)$ counts the number of monomer-dimer configurations with k dimers. The weight parameter λ reflects the contribution of a dimer to the energy of the system.

Our main goal in this section is the development of an algorithm for approximating Z_G at an arbitrary point $\lambda \geq 0$. The running time of the algorithm is $\mathrm{poly}(n, \epsilon, \max\{\lambda, 1\})$, where ϵ, as usual, controls the relative error that will be tolerated in the output. Thus the algorithm will meet the specification of an FPRAS for Z_G, provided λ is specified in unary notation. Our approach is to simulate a suitable Markov chain $\mathcal{M}_{\mathrm{match}}(\lambda)$, parameterised on the the graph G and edge weight λ. The state space, Ω, is the set of all matchings in G, and the transitions are constructed so that the chain is ergodic with stationary distribution π_λ given by

$$\pi_\lambda(M) = \frac{\lambda^{|M|}}{Z(\lambda)}. \tag{3.2}$$

(Since G is fixed from now on, we drop the subscript from Z.) In other words, the stationary probability of each matching (monomer-dimer configuration) is proportional to its weight in the partition function (3.1). The Markov chain $\mathcal{M}_{\mathrm{match}}(\lambda)$, if simulated for sufficiently many steps, provides a method of sampling matchings from the distribution π_λ.

It is not hard to construct a Markov chain $\mathcal{M}_{\mathrm{match}}(\lambda)$ with the right asymptotic properties. Consider the chain in which transitions from any matching M are made according to the following rule:

1. with probability $\frac{1}{2}$ let $M' = M$; otherwise,

2. select an edge $e = \{u, v\} \in E$ u.a.r. and set

$$M' = \begin{cases} M - e & \text{if } e \in M; \\ M + e & \text{if both } u \text{ and } v \text{ are unmatched in } M; \\ M + e - e' & \text{if exactly one of } u \text{ and } v \text{ is matched in } M \\ & \quad\quad \text{and } e' \text{ is the matching edge;} \\ M & \text{otherwise;} \end{cases}$$

3. go to M' with probability $\min\{1, \pi_\lambda(M')/\pi_\lambda(M)\}$.

It is helpful to view this chain as follows. There is an underlying graph defined on the set of matchings Ω in which the neighbours of matching M are all matchings M' that differ from M via one of the following local perturbations: an edge is removed from M (a $\downarrow$-*transition*); an edge is added to M (a $\uparrow$-*transition*); or a new edge is exchanged with an edge in M (a $\leftrightarrow$-*transition*). Transitions from M are made by first selecting a neighbour M' u.a.r., and then actually making, or *accepting* the transition with probability $\max\{1, \pi_\lambda(M')/\pi_\lambda(M)\}$. Note that the ratio appearing in this expression is easy to compute: it is just λ^{-1}, λ or 1 respectively, according to the type of the transition.

As the reader may easily verify, this acceptance probability is constructed so that the transition probabilities $P(M, M')$ satisfy the detailed balance condition

$$Q(M, M') = \pi_\lambda(M)P(M, M') = \pi_\lambda(M')P(M', M), \quad \text{for all } M, M' \in \Omega,$$

i.e., $\mathcal{M}_{\mathrm{match}}(\lambda)$ is reversible. This fact, together with the observation that $\mathcal{M}_{\mathrm{match}}(\lambda)$ is irreducible (i.e., all states communicate, for example via the empty matching) and aperiodic (by step 1, the self-loop probabilities $P(M, M)$ are all non-zero), ensures that $\mathcal{M}_{\mathrm{match}}(\lambda)$ is ergodic with stationary distribution π_λ, as required. The device of performing random walk on a connected graph with acceptance probabilities of this form is well known in Monte Carlo physics under the name of the "Metropolis process". Clearly, it can be used to achieve any desired stationary distribution π for which the ratio $\pi(u)/\pi(v)$ for neighbours u, v can be computed easily. It is also the standard mechanism used in combinatorial optimisation by simulated annealing.

Having constructed a family of Markov chains with stationary distribution π_λ, our next task is to explain how samples from this distribution can be used to obtain a reliable statistical estimate of $Z(\lambda)$ at a specified point $\lambda = \widehat{\lambda} \geq 0$. Our strategy is to express $Z(\widehat{\lambda})$ as the product

$$Z(\widehat{\lambda}) = \frac{Z(\lambda_r)}{Z(\lambda_{r-1})} \times \frac{Z(\lambda_{r-1})}{Z(\lambda_{r-2})} \times \cdots \frac{Z(\lambda_2)}{Z(\lambda_1)} \times \frac{Z(\lambda_1)}{Z(\lambda_0)} \times Z(\lambda_0), \tag{3.3}$$

where $0 = \lambda_0 < \lambda_1 < \lambda_2 < \cdots < \lambda_{r-1} < \lambda_r = \widehat{\lambda}$ is a suitably chosen sequence of values. Note that $Z(\lambda_0) = Z(0) = 1$. We will then estimate each factor $Z(\lambda_i)/Z(\lambda_{i-1})$ in this product by sampling from the distribution π_{λ_i}. This approach is analogous to that used in the context of independent sets in the proof of Theorem 1.2.1; refer in particular to equation (1.4). For reasons that will become clear shortly, we will use the sequence of values $\lambda_1 = (2|E|)^{-1}$ and $\lambda_i = (1 + \frac{1}{n})^{i-1}\lambda_1$ for $1 \le i < r$. The length r of the sequence is taken to be minimal such that $(1 + \frac{1}{n})^{r-1}\lambda_1 \ge \widehat{\lambda}$, so we have the bound

$$r \le \left\lceil 2n\big(\ln \widehat{\lambda} + \ln(2|E|)\big)\right\rceil + 1. \tag{3.4}$$

To estimate the ratio $Z(\lambda_i)/Z(\lambda_{i-1})$, we will express it, or rather its reciprocal, as the expectation of a suitable random variable. Specifically, define the random variable $Z_i(M) = \big(\frac{\lambda_{i-1}}{\lambda_i}\big)^{|M|}$, where M is a matching chosen from the distribution π_{λ_i}. Then we have

$$\mathbf{E}(Z_i) = \sum_M \left(\frac{\lambda_{i-1}}{\lambda_i}\right)^{|M|} \frac{\lambda_i^{|M|}}{Z(\lambda_i)} = \frac{1}{Z(\lambda_i)} \sum_M \lambda_{i-1}^{|M|} = \frac{Z(\lambda_{i-1})}{Z(\lambda_i)}.$$

Thus the ratio $\rho_i = Z(\lambda_{i-1})/Z(\lambda_i)$ can be estimated by sampling matchings from the distribution π_{λ_i} and computing the sample mean of Z_i. Following (3.3), our estimator of $Z(\widehat{\lambda})$ will be the product of the reciprocals of these estimated ratios. Summarising this discussion, our algorithm can be written down as follows:

Step 1 Compute the sequence $\lambda_1 = (2|E|)^{-1}$ and $\lambda_i = \big(1 + \frac{1}{n}\big)^{i-1}\lambda_1$ for $1 \le i < r$, where r is the least integer such that $\big(1 + \frac{1}{n}\big)^{r-1}\lambda_1 \ge \widehat{\lambda}$. Set $\lambda_0 = 0$ and $\lambda_r = \widehat{\lambda}$.

Step 2 For each value $\lambda = \lambda_1, \lambda_2, \ldots, \lambda_r$ in turn, compute an estimate X_i of the ratio ρ_i as follows:

- by performing S independent simulations of the Markov chain $\mathcal{M}_{\text{match}}(\lambda_i)$, each of length T_i, obtain an independent sample of size S from (close to) the distribution π_{λ_i};
- let X_i be the sample mean of the quantity $\big(\frac{\lambda_{i-1}}{\lambda_i}\big)^{|M|}$.

Step 3 Output the product $Y = \prod_{i=1}^{r} X_i^{-1}$.

Figure 3.1: Algorithm MatchSample

To complete the description of the algorithm, we need to specify the sample size S in Step 2, and the number of simulation steps T_i required for each sample. Our goal is to show that, with suitable values for these quantities, Algorithm MatchSample is an FPRAS for $Z(\lambda)$.

The issue of the sample size S is straightforward. Now $\mathrm{e}^{-1} \le Z_i \le 1$ and so using Lemma 1.2.1 of Chapter 1 we see

Proposition 3.1.1 *In Algorithm* MATCHSAMPLE, *suppose the sample size S in Step 2 is $S = \lceil 17e^2\epsilon^{-2}r \rceil$, and that the simulation length T_i is large enough that the variation distance of $\mathcal{M}_{\text{match}}(\lambda_i)$ from its stationary distribution π_{λ_i} is at most $\epsilon/(3er)$. Then the output random variable Y satisfies*

$$\mathbf{Pr}\big((1-\epsilon)Z(\widehat{\lambda}) \le Y \le (1+\epsilon)Z(\widehat{\lambda})\big) \ge \tfrac{3}{4}.$$

Since r is a relatively small quantity (essentially linear in n: see (3.4)), this result means that a modest sample size at each stage suffices to ensure a good final estimate Y, provided of course that the samples come from a distribution that is close enough to π_{λ_i}.

It is in determining the number of simulation steps, T_i, required to achieve this that the meat of the analysis lies: of course, this is tantamount to investigating the mixing time of the Markov chain $\mathcal{M}_{\text{match}}(\lambda_i)$. Our main task in this section will be to show:

Proposition 3.1.2 *The mixing time of the Markov chain $\mathcal{M}_{\text{match}}(\lambda)$ satisfies*

$$\tau_X(\epsilon) \le 4|E|n\lambda'\big(n(\ln n + \ln \lambda') + \ln \epsilon^{-1}\big),$$

where $\lambda' = \max\{\lambda, 1\}$.

The proof of this result will make use of the full power of the machinery introduced in Section 2.2.3 of Chapter 2. Note that Proposition 3.1.2 is a very strong statement: it says that we can sample from (close to) the complex distribution π_λ over the exponentially large space of matchings in G, by performing a Markov chain simulation of length only a low-degree polynomial in the size of G.[1]

According to Proposition 3.1.1, we require a variation distance of $\epsilon/(3er)$, so Proposition 3.1.2 tells us that it suffices to take

$$T_i = \big\lceil 4|E|n\lambda_i'\big(n(\ln n + \ln \lambda_i') + \ln(3er/\epsilon)\big)\big\rceil. \tag{3.5}$$

This concludes our specification of the Algorithm MATCHSAMPLE.

Before proceeding to prove the above statements, let us convince ourselves that together they imply that Algorithm MATCHSAMPLE is an FPRAS for $Z(\lambda)$. First of all, Proposition 3.1.1 ensures that the output of Algorithm MATCHSAMPLE satisfies the requirements of an FPRAS for Z. It remains only to verify that the running time is bounded by a polynomial in n, $\widehat{\lambda}'$ and ϵ^{-1}. Evidently the running time is dominated by the number of Markov chain simulations steps, which is $\sum_{i=1}^{r} ST_i$; since T_i increases with i, this is at most rST_r. Substituting the upper bound for r from (3.4), and values

[1] Incidentally, we should point out that Proposition 3.1.2 immediately tells us that we can sample monomer-dimer configurations from the canonical distribution π_λ, in time polynomial in n and λ'. This is in itself an interesting result, and allows estimation of the expectation of many quantities associated with monomer-dimer configurations.

for S from Proposition 3.1.1 and T_r from (3.5), we see that the overall running time of Algorithm MATCHSAMPLE is bounded by[2]

$$O\big(n^4|E|\widehat{\lambda}'(\ln n\widehat{\lambda}')^3\epsilon^{-2}\big),$$

which grows only polynomially with n, $\widehat{\lambda}'$ and ϵ^{-1}. We have therefore proved

Theorem 3.1.1 *Algorithm* MATCHSAMPLE *is an FPRAS for the partition function of an arbitrary monomer-dimer system.*

We turn now to the question of proving Proposition 3.1.2. Our strategy will be to carefully choose a collection of canonical paths $\Gamma = \{\gamma_{XY} : X, Y \in \Omega\}$ in the Markov chain $\mathcal{M}_{\mathrm{match}}(\lambda)$ for which the "bottleneck" measure $\bar{\rho}(\Gamma)$ of Section 2.2.3 is small. We can then appeal to Theorem 2.2.4 and Corollary 2.1.1 to bound the mixing time. Specifically, we shall show that our paths satisfy

$$\bar{\rho}(\Gamma) \leq 4|E|n\lambda'. \tag{3.6}$$

Since the number of matchings in G is certainly bounded above by $(2n)!$, the stationary probability $\pi_\lambda(X)$ of any matching X is bounded below by $\pi_\lambda(X) \geq 1/(2n)!\lambda'^n$. Using (3.6) and the fact that $\ln n! \leq n\ln n$, the bound on the mixing time in Proposition 3.1.2 can now be read off from Theorem 2.2.4 and Corollary 2.1.1.

It remains for us to find a set of canonical paths Γ satisfying (3.6). For a pair of matchings X, Y in G, we define the canonical path γ_{XY} as follows. Consider the symmetric difference $X \oplus Y$. A moment's reflection should convince the reader that this consists of a disjoint collection of paths in G (some of which may be closed cycles), each of which has edges that belong alternately to X and to Y. Now suppose that we have fixed some arbitrary ordering on all simple paths in G, and designated in each of them a so-called "start vertex", which is arbitrary if the path is a closed cycle but must be an endpoint otherwise. This ordering induces a unique ordering $P_1, P_2, \dots, P_m$ on the paths appearing in $X \oplus Y$. The canonical path from X to Y involves "unwinding" each of the P_i in turn as follows. There are two cases to consider:

1. *P_i is not a cycle.* Let P_i consist of the sequence $(v_0, v_1, \dots, v_l)$ of vertices, with v_0 the start vertex. If $(v_0, v_1) \in Y$, perform a sequence of $\leftrightarrow$-transitions replacing (v_{2j+1}, v_{2j+2}) by (v_{2j}, v_{2j+1}) for $j = 0, 1, \dots$, and finish with a single $\uparrow$-transition if l is odd. If on the other hand $(v_0, v_1) \in X$, begin with a $\downarrow$-transition removing (v_0, v_1) and proceed as before for the reduced path $(v_1, \dots, v_l)$.

[2] In deriving the O-expression, we have assumed w.l.o.g. that $T_r = O\big(|E|n^2\widehat{\lambda}'\ln n\widehat{\lambda}'\big)$. This follows from (3.5) with the additional assumption that $\ln\epsilon^{-1} = O(n\ln n)$. This latter assumption is justified since the problem can always be solved exactly by exhaustive enumeration in time $O(n(2n)!)$, which is $O(\epsilon^{-2})$ if $\ln\epsilon^{-1}$ exceeds the above bound.

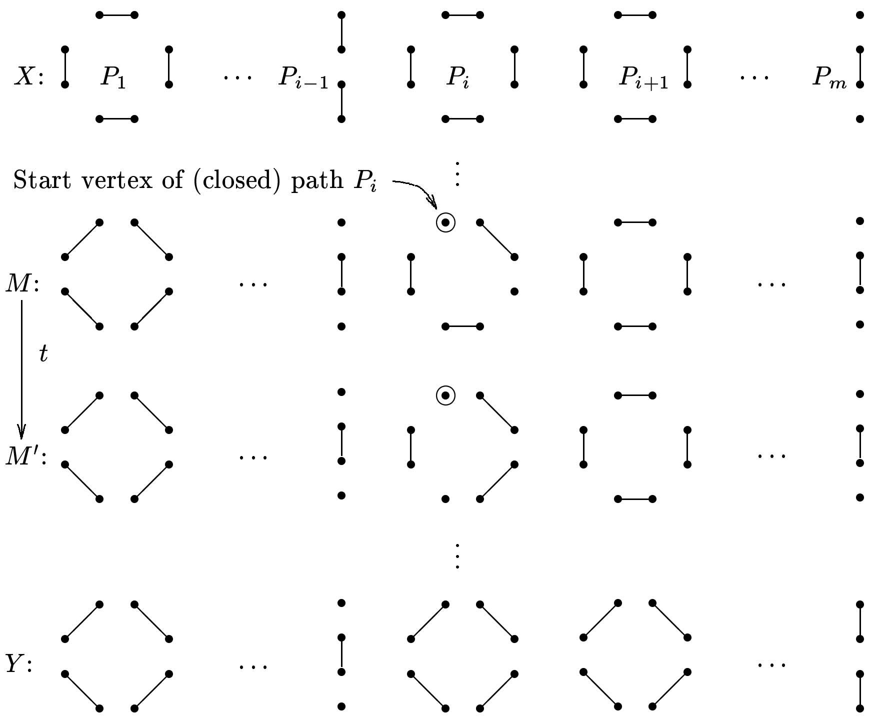

Figure 3.2: A transition t in the canonical path from X to Y

2. *P_i is a cycle.* Let P_i consist of the sequence $(v_0, v_1, \dots, v_{2l+1})$ of vertices, where $l \geq 1$, v_0 is the start vertex, and $(v_{2j}, v_{2j+1}) \in X$ for $0 \leq j \leq l$, the remaining edges belonging to Y. Then the unwinding begins with a $\downarrow$-transition to remove (v_0, v_1). We are left with an open path O with endpoints v_0, v_1, one of which must be the start vertex of O. Suppose v_k, $k \in \{0, 1\}$, is *not* the start vertex. Then we unwind O as in (i) above but treating v_k as the start vertex. This trick serves to distinguish paths from cycles, as will prove convenient shortly.

This concludes our definition of the family of canonical paths Γ. Figure 3.2 will help the reader picture a typical transition t on a canonical path from X to Y. The path P_i (which happens to be a cycle) is the one currently being unwound; the paths $P_1, \dots, P_{i-1}$ to the left have already been processed, while the ones $P_{i+1}, \dots, P_m$ are yet to be dealt with.

We now proceed to bound the "bottleneck" measure $\bar{\rho}(\Gamma)$ for these paths. Let ϵ be an arbitrary edge in the Markov chain, i.e., a transition from M to $M' \neq M$, and let $\mathrm{cp}(\epsilon) = \{(X, Y) : \gamma_{XY} \ni \epsilon\}$ denote the set of all canonical paths that use ϵ. (We use the notation ϵ in place of e here to avoid confusion with edges of G.) We will obtain a bound on the total weight of all paths that pass through ϵ by defining an injective mapping $\eta_\epsilon : \mathrm{cp}(\epsilon) \to \Omega$. What we would like to do is to set $\eta_\epsilon(X, Y) = X \oplus Y \oplus (M \cup M')$; the intuition for this is that $\eta_\epsilon(X, Y)$ should agree with X on paths that have already

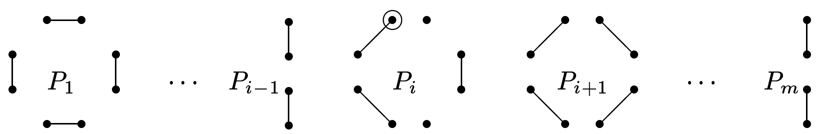

Figure 3.3: The corresponding encoding $\eta_t(X, Y)$

been unwound, and with Y on paths that have not yet been unwound. However, there is a minor complication concerning the path that we are currently processing: in order to ensure that $\eta_\epsilon(X, Y)$ is indeed a matching, we may — as we shall see — have to remove from it the edge of X adjacent to the start vertex of the path currently being unwound: we shall call this edge e_{XYt}. This leads us to the following definition of the mapping η_ϵ:

$$\eta_\epsilon(X,Y) = \begin{cases} X \oplus Y \oplus (M \cup M') - e_{XYt}, & \text{if } \epsilon \text{ is a } \leftrightarrow\text{-transition} \\ & \text{and the current path is a cycle;} \\ X \oplus Y \oplus (M \cup M'), & \text{otherwise.} \end{cases}$$

Figure 3.5 illustrates the encoding $\eta_t(X, Y)$ that would result from the transition t on the canonical path sketched in Figure 3.2.

Let us check that $\eta_\epsilon(X, Y)$ is always a matching. To see this, consider the set of edges $A = X \oplus Y \oplus (M \cup M')$, and suppose that some vertex, u say, has degree two in A. (Since $A \subseteq X \cup Y$, no vertex degree can exceed two.) Then A contains edges $\{u, v_1\}, \{u, v_2\}$ for distinct vertices v_1, v_2, and since $A \subseteq X \cup Y$, one of these edges must belong to X and the other to Y. Hence both edges belong to $X \oplus Y$, which means that neither can belong to $M \cup M'$. Following the form of $M \cup M'$ along the canonical path, however, it is clear that there can be at most one such vertex u; moreover, this happens precisely when the current path is a cycle, u is its start vertex, and ϵ is a $\leftrightarrow$-transition. Our definition of η_ϵ removes one of the edges adjacent to u in this case, so all vertices in $\eta_\epsilon(X, Y)$ have degree at most one, i.e., $\eta_\epsilon(X, Y)$ is indeed a matching.

We now have to check that η_ϵ is injective. It is immediate from the definition of η_ϵ that the symmetric difference $X \oplus Y$ can be recovered from $\eta_\epsilon(X, Y)$ using the relation

$$X \oplus Y = \begin{cases} \eta_\epsilon(X,Y) \oplus (M \cup M') + e_{XYt}, & \text{if } \epsilon \text{ is a } \leftrightarrow\text{-transition} \\ & \text{and the current path is a cycle;} \\ \eta_\epsilon(X,Y) \oplus (M \cup M'), & \text{otherwise.} \end{cases}$$

Note that, once we have formed the set $\eta_\epsilon(X, Y) \oplus (M \cup M')$, it will be apparent whether the current path is a cycle from the sense of unwinding. (Note that e_{XYt} is the unique edge that forms a cycle when added to the path.) Given $X \oplus Y$, we can at once infer the sequence of paths $P_1, P_2, \ldots, P_m$ that have to be unwound along the canonical path from X to Y, and the transition t tells us which of these, P_i say, is the path currently being unwound. The partition of $X \oplus Y$ into X and Y is now straightforward: X has the same parity as $\eta_\epsilon(X, Y)$ on paths $P_1, \ldots, P_{i-1}$, and the same parity as M on paths

$P_{i+1}, \ldots, P_m$. Finally, the reconstruction of X and Y is completed by noting that $X \cap Y = M - (X \oplus Y)$, which is immediate from the definition of the paths. Hence X and Y can be uniquely recovered from $\eta_\epsilon(X,Y)$, so η_ϵ is injective.

We are almost done. What we now require in addition is that η_ϵ be "weight-preserving," in the sense that $Q(\epsilon)\pi_\lambda(\eta_\epsilon(X,Y)) \approx \pi_\lambda(X)\pi_\lambda(Y)$. More precisely, we will show in a moment that

$$\pi_\lambda(X)\pi_\lambda(Y) \leq 2|E|\lambda'^2 Q(\epsilon)\pi_\lambda(\eta_\epsilon(X,Y)). \tag{3.7}$$

First, let us see why we need a bound of this form in order to estimate $\bar{\rho}$. We have

$$\begin{aligned}
\frac{1}{Q(\epsilon)} \sum_{\gamma_{XY} \ni \epsilon} \pi_\lambda(X)\pi_\lambda(Y)|\gamma_{XY}| &\leq 2|E|\lambda'^2 \sum_{\gamma_{XY} \ni \epsilon} \pi_\lambda(\eta_\epsilon(X,Y))\,|\gamma_{XY}| \\
&\leq 4|E|n\lambda'^2 \sum_{\gamma_{XY} \ni \epsilon} \pi_\lambda(\eta_\epsilon(X,Y)) \\
&\leq 4|E|n\lambda'^2, \qquad (3.8)
\end{aligned}$$

where the second inequality follows from the fact that the length of any canonical path is bounded by $2n$, and the last inequality from the facts that η_ϵ is injective and π_λ is a probability distribution.

It remains for us to prove inequality (3.7). Before we do so, it is helpful to notice that $Q(\epsilon) = (2|E|)^{-1}\min\{\pi_\lambda(M), \pi_\lambda(M')\}$, as may easily be verified from the definition of $\mathcal{M}_{\mathrm{match}}(\lambda)$. We now distinguish four cases:

1. *ϵ is a $\downarrow$-transition.* Suppose $M' = M - e$. Then $\eta_\epsilon(X,Y) = X \oplus Y \oplus M$, so, viewed as multisets, $M \cup \eta_\epsilon(X,Y)$ and $X \cup Y$ are identical. Hence we have

$$\begin{aligned}
\pi_\lambda(X)\pi_\lambda(Y) &= \pi_\lambda(M)\pi_\lambda(\eta_\epsilon(X,Y)) \\
&= \frac{2|E|Q(\epsilon)}{\min\{\pi_\lambda(M), \pi_\lambda(M')\}} \times \pi_\lambda(M)\pi_\lambda(\eta_\epsilon(X,Y)) \\
&= 2|E|Q(\epsilon)\max\{1, \pi_\lambda(M)/\pi_\lambda(M')\}\pi_\lambda(M)\pi_\lambda(\eta_\epsilon(X,Y)) \\
&\leq 2|E|\lambda' Q(\epsilon)\pi_\lambda(\eta_\epsilon(X,Y)),
\end{aligned}$$

   from which (3.7) follows.

2. *ϵ is a $\uparrow$-transition.* This is handled by a symmetrical argument to (i) above, with the roles of M and M' interchanged.

3. *ϵ is a $\leftrightarrow$-transition and the current path is a cycle.* Suppose $M' = M + e - e'$, and consider the multiset $M \cup \eta_\epsilon(X,Y)$. Then $\eta_\epsilon(X,Y) = X \oplus Y \oplus (M+e) - e_{XYt}$, so the multiset $M \cup \eta_\epsilon(X,Y)$ differs from $X \cup Y$ only in that e and e_{XYt} are missing from it. Thus we have

$$\begin{aligned}
\pi_\lambda(X)\pi_\lambda(Y) &\leq \lambda'^2 \pi_\lambda(M)\pi_\lambda(\eta_\epsilon(X,Y)) \\
&= 2|E|\lambda'^2 Q(\epsilon)\pi_\lambda(\eta_\epsilon(X,Y)),
\end{aligned}$$

since in this case $\pi_\lambda(M) = \pi_\lambda(M')$, and so $Q(\epsilon) = (2|E|)^{-1}\pi_\lambda(M)$. Thus (3.7) is again satisfied.

4. *ϵ is a $\leftrightarrow$-transition and the current path is not a cycle.* This is identical with (iii) above, except that the edge e_{XYt} does not appear in the analysis. Accordingly, the bound is
$$\pi_\lambda(X)\pi_\lambda(Y) \leq 2|E|\lambda' Q(\epsilon)\pi_\lambda(\eta_\epsilon(X,Y)).$$

This concludes our proof of (3.7). We may now deduce from (3.8), that $\bar{\rho}(\Gamma) \leq 4|E|n\lambda'^2$. However, one additional observation will allow us to improve the bound to $\bar{\rho}(\Gamma) \leq 4|E|n\lambda'$, which is what we claimed in (3.6). Looking at the above case analysis we see that, in all cases except case (iii), (3.7), and hence (3.8), actually hold with λ'^2 replaced by λ'. But in case (iii) we can argue that $\eta_\epsilon(X,Y)$ must have such a restricted form that $\sum_{\gamma_{XY}\ni\epsilon} \pi_\lambda(\eta_\epsilon(X,Y))$ is bounded above by λ'^{-1}. Using this fact in the final inequality in (3.8), we get the improved upper bound of $4|E|n\lambda'$ in this case, and hence in all cases. This will complete our verification of the bound (3.6) on $\bar{\rho}(\Gamma)$.

To justify the above claim, note that $\eta_\epsilon(X,Y)$ has at least two unmatched vertices, namely the start vertex of the current cycle and the vertex that is common to both e and e'. Moreover, in $\eta_\epsilon(X,Y) \oplus M$ these vertices are linked by an alternating path that starts and ends with an edge of M. So we may associate with each matching $\eta_\epsilon(X,Y)$ another matching, say $\eta'_\epsilon(X,Y)$, obtained by augmenting $\eta_\epsilon(X,Y)$ along this path. But this operation is uniquely reversible, so all matchings $\eta'_\epsilon(X,Y)$ created in this way are distinct. Moreover, $\pi_\lambda(\eta_\epsilon(X,Y)) = \lambda\pi_\lambda(\eta_\epsilon(X,Y))$. Hence we have $\sum \pi_\lambda(\eta_\epsilon(X,Y)) = \lambda^{-1}\sum \pi_\lambda(\eta'_\epsilon(X,Y)) \leq \lambda^{-1}$, so $\sum \pi_\lambda(\eta_\epsilon(X,Y)) \leq \lambda'^{-1}$ as claimed.

3.2 Perfect Matchings

The question of whether there exists an FPRAS for the permanent of an arbitrary 0,1-matrix has recently been positivley resolved by Jerrum, Sinclair and Vigoda [**?**]. Thus there is an FPRAS for the number of perfect matchings in an arbitrary bipartite graph. Their result does not seem to carry over for arbitrary graphs and so we first concentrate on seeing how to use the methods and results of the previous section to construct an FPRAS that covers many cases, even a majority in some sense. To state the result precisely, we will use the perfect matching formulation. Let $G = (V,E)$ be a graph with $|V| = 2n$. A special role will be played in the result by the number of *near-perfect* matchings in G, i.e., matchings with exactly two unmatched vertices. Following the notation of the previous section, let us write $m_k = m_k(G)$ for the number of k-matchings in G. Then the number of perfect matchings is m_n, and the number of near-perfect matchings is m_{n-1}.

Theorem 3.2.1 *There exists a randomized approximation scheme for the number of perfect matchings m_n whose running time is polynomial in n, ϵ^{-1} and the ratio m_{n-1}/m_n.*

Note that this algorithm is not in general an FPRAS, since there exist $2n$-vertex graphs G for which the ratio m_{n-1}/m_n is exponential in n. However, it turns out that these examples are atypical in the sense that the probability that a randomly selected G on $2n$ vertices violates the inequality $m_{n-1}/m_n \leq 4n$ tends to 0 as $n \to \infty$. Thus the above algorithm constitutes an FPRAS for almost all graphs; moreover, the condition that the ratio m_{n-1}/m_n be bounded by a specified polynomial in n can be tested for an arbitrary graph in polynomial time. It is also known that *every* sufficiently dense graph (specifically, those in which every vertex has degree at least $\frac{1}{2}n$) satisfies $m_{n-1}/m_n = \mathrm{O}(n^2)$. Moreover, it has been shown ratio m_{n-1}/m_n is guaranteed to be small for a wide class of homogeneous graphs G, including the important case of geometric lattice graphs in any number of dimensions.

Our approximation algorithm for the number of perfect matchings follows quite painlessly from our results about the matchings problem derived in the previous section. Note that m_n is precisely the leading coefficient of the partition function $Z_G(\lambda)$ of the monomer-dimer system associated with G (see (3.1)). In the previous section, we saw how to sample matchings in G from the distribution

$$\pi_\lambda(M) = \frac{\lambda^{|M|}}{Z_G(\lambda)} = \frac{\lambda^{|M|}}{\sum_{k=0}^{n} m_k \lambda^k} \tag{3.9}$$

for any desired $\lambda > 0$, in time polynomial in n and $\lambda' = \max\{\lambda, 1\}$, by Monte Carlo simulation of the Markov chain $\mathcal{M}_{\mathrm{match}}(\lambda)$. We also saw how this fact can be used to compute $Z_G(\lambda)$ to good accuracy in time polynomial in n and λ'. Suppose then that we have computed a good estimate $\widehat{Z}_G(\lambda)$ of $Z_G(\lambda)$. Then we can get a good estimator for m_n by sampling matchings from the distribution π_λ and computing the proportion, X, of the sample that are perfect matchings; since $\mathbf{E}X = m_n\lambda^n/Z_G(\lambda)$, our estimator is $Y = X\lambda^{-n}\widehat{Z}_G(\lambda)$.

The sample size required to ensure a good estimate depends on the variance of a single sample, or more precisely on the quantity $(\mathbf{E}X)^{-1}$. Clearly, by making λ large enough, we can make this quantity, and hence the sample size, small: this corresponds to placing very large weight on the perfect matchings, so that their proportion can be estimated well by random sampling. How large does λ have to be? This analysis is eased by a beautiful fact. A sequence $a_1, a_2, \ldots, a_n$ of positive reals is *log-concave* if $a_{k-1}a_{k+1} \leq a_k^2$ for $k = 1, 2, \ldots, n-1$.

Lemma 3.2.1 *The sequence $m_0, m_1, \ldots, m_n$ is log-concave.*

Proof Let $M_k = M_k(G)$ be the set of k-matchings of G. Thus $m_k = |M_k(G)|$. We need to show that $m_{k-1}m_{k+1} \leq m_k^2$ and so we can assume that $m_{k+1} > 0$. Let $A = M_{k+1} \times M_{k-1}$ and $B = M_k \times M_k$.

If M, M' are matchings then we know that $M \oplus M$ consists of paths and cycles. Let a path of $M \oplus M$ be an M-path if it contains more M-edges than M'-edges and an M' path if the reverse is true. For any pair $(M, M') \in A$ the number of M paths exceeds the number of M' paths by exactly two. We partition A into disjoint classes $A_r, r = 1, 2, \ldots, k$ where

$$A_r = \{(M, M') \in A : \ M \oplus M' \text{ contains } r+1 \ M-\text{paths and } r-1 \ M'-\text{paths}\}.$$

Similarly the sets

$$B_r = \{(M, M') \in B : \ M \oplus M' \text{ contains } r \ M-\text{paths and } r \ M'-\text{paths}\}.$$

partition B. The lemma will follow from the fact that $|A_r| \leq |B_r|$ for each $r > 0$.

Let us call a pair $(L, L') \in B_r$ *reachable* from $(M, M') \in A_r$ iff $L \oplus L' = M \oplus M'$ and $L = M \oplus P$ for some M-path P of $M \oplus M'$. Clearly the number of elements of B_r reachable from a given $(M, M') \in A_r$ is $r + 1$. Conversely, any given element of B_r is reachable from precisely r elements of A_r. Hence if $|A_r| > 0$ we have $|B_r|/|A_r| = (r+1)/r > 1$. □

As a consequence, it follows that that $m_k/m_n \leq (m_{n-1}/m_n)^{n-k}$. This means that, if we take $\lambda \geq m_{n-1}/m_n$, we get

$$\mathbf{E}X = \frac{m_n \lambda^n}{Z_G(\lambda)} = \frac{m_n \lambda^n}{\sum_{k=0}^{n} m_k \lambda^k} \geq \frac{1}{n+1}, \tag{3.10}$$

which implies that the sample size required grows only linearly with n. Thus it is enough to take λ about as large as the ratio m_{n-1}/m_n. Of course we do not have a priori knowledge of this ratio and so we run the algorithm with $\lambda = 2, 4, 8, \ldots$ until we find that at least $1/n$ proportion of the matchings produced are perfect. Since the time required to generate a single sample grows linearly with λ (see Proposition 3.1.2), the running time of the overall algorithm is polynomial in n, ϵ^{-1} and the ratio m_{n-1}/m_n, as claimed.

We conclude thsi section by mentioning some extensions. First of all, it is not hard to see, again using the log-concavity property, that the above technique can be extended to approximate the entire sequence (m_k), or equivalently all the coefficients of the monomer-dimer partition function. The running time per coefficient is no worse than for m_n.

A Special Case: Vertex Transitive graphs

In this section we discuss a class of graphs for which we can prove that m_{n-1}/m_n is polynomially bounded. A grpah G is *vertex transitive* if for every pair of vertices u, v there is an automorphism g of G such that $g(u) = v$. As an example consider the *discrete torus* $T_{d,L}$ in d dimensions. Here we take $V = \{0, 1, \ldots, L-1\}$ for some $L > 0$ and two vertices x, y are adjacent if there exists an index j such that $x_i = y_i$, $i \neq j$ and $|x_j - y_j| = 1 \bmod L$.

Theorem 3.2.2 *In any vertex transitive graph G, the number of near-perfect matchings in G exceeds the number of perfect matchings by a factor at most n^3.*

Proof Let G be any graph with transitive automorphism group. Denote by $\mathcal{M}$ the set of all perfect matchings in G, by $\mathcal{N}(u, v)$ the set of near-perfect matchings that leave vertices u and v uncovered, and by $\mathcal{N} = \bigcup_{u,v} \mathcal{N}(u, v)$ the set of all near-perfect matchings. Let $\mu = |\mathcal{M}|$, $\nu = \max_{u,v} |\mathcal{N}(u, v)|$, and select vertices u_0 and v_0 satisfying $|\mathcal{N}(u_0, v_0)| = \nu$. We assume, contrary to the statement of the theorem, that $\mu < \nu/2n$, and obtain a contradiction.

Let u, v be any pair of non-adjacent vertices. We show that there exists a vertex u' with $\operatorname{dist}(u', v) < \operatorname{dist}(u, v)$ satisfying $|\mathcal{N}(u', v)| \geq |\mathcal{N}(u, v)| - \nu/2n$. By induction on distance (starting with the base case $|\mathcal{N}(u_0, v_0)| = \nu$) it follows that there exists a pair of adjacent vertices (u, v) satisfying $|\mathcal{N}(u, v)| \geq \nu/2$ and hence that $\mu = |\mathcal{M}| \geq \nu/2$, contradicting our initial assumption.

So suppose u and v are non-adjacent, and let u' be any vertex adjacent to u satisfying $\operatorname{dist}(u', v) < \operatorname{dist}(u, v)$. Let g be any automorphism of G mapping u_0 to u' and let v' be the image of v_0 under g. If $v' = v$ we are done, since then $|\mathcal{N}(u', v)| = |\mathcal{N}(u_0, v_0)| = \nu$. So assume the contrary. Define a mapping

$$f : \mathcal{N}(u, v) \times \mathcal{N}(u', v') \to \mathcal{M} \times \mathcal{N}(v, v') \cup \mathcal{N}(u, v') \times \mathcal{N}(u', v)$$

as follows. Let $N \in \mathcal{N}(u, v)$ be a "red" and $N' \in \mathcal{N}(u', v')$ a "blue" near-perfect matching. $N \oplus N'$ consists of two alternating paths together with a number of cycles. We distinguish two cases.

1. There is a blue-blue path from u to v and red-red path from u' to v', *or* there is a blue-red path from u to v' and a red-blue path from u' to v. In this case add a red edge $\{u, u'\}$ and exchange the colours along the (u', v') or (u', v) path, as appropriate. This operation yields a perfect matching in $\mathcal{M}$, and a near-perfect matching in $\mathcal{N}(v, v')$.

2. There is a blue-red path from u to u' and a blue-red path from v to v'. In this case, exchange the colours along the (u, u') path to yield a blue near-perfect matching in $\mathcal{N}(u, v')$ and a red near-perfect matching in $\mathcal{N}(u', v)$.

The operations described above are reversible, so the mapping f is injective. Thus

$$|\mathcal{N}(u, v)| \times |\mathcal{N}(u', v')| \leq |\mathcal{M}| \times |\mathcal{N}(v, v')| + |\mathcal{N}(u, v')| \times |\mathcal{N}(u', v)|.$$

Since $|\mathcal{N}(u', v')| = |\mathcal{N}(u_0, v_0)| = \nu$ and $|\mathcal{M}| = \mu < \nu/2n$ by assumption, the required inequality $|\mathcal{N}(u', v)| \geq |\mathcal{N}(u, v)| - \nu/2n$ follows. □

3.3 The Permanent

Let $G = (V_1, V_2, E)$ be a bipartite graph on $n+n$ vertices. Let $\mathcal{M}_k = \mathcal{M}_k(G)$ be the set of k-matchings of G. There is a rapidly mixing Markov chain on state space $\Omega = \mathcal{M}_{n-1} \cup \mathcal{M}_n$ similar to that described in Section 3.1 that has a uniform steady state distribution. Of course if m_{n-1}/m_n is too large then it will tend to only generate near-perfect matchings within a reasonable time limit. The idea from [?] is to modify this chain so that the steady state is still uniform over perfect matchings and the near-perfect matchings have sufficiently smaller weight.

Theorem 3.3.1 *There exists a fully-polynomial randomized approximation scheme for the permanent of an arbitrary $n \times n$ 0-1 matrix A.*

It will actually prove technically convenient to introduce edge weights also. Thus for each edge $(y, z) \in E$, we introduce a positive weight $\lambda(y, z)$, which we call its *activity.* We extend the notion of activities to matchings M (of any cardinality) by $\lambda(M) = \prod_{(i,j)\in M} \lambda(i, j)$. Similarly, for a set of matchings $\mathcal{S}$ we define $\lambda(\mathcal{S}) = \sum_{M\in\mathcal{S}} \lambda(M)$. For our purposes, the advantage of edge weights is that they allow us to work with the complete graph $K_{n,n}$ on $n + n$ vertices, rather than with an arbitrary graph $G = (V_1, V_2, E)$: we can do this by setting $\lambda(e) = 1$ for $e \in E$, and $\lambda(e) \leq 1/n!$ for $e \notin E$. This ensures that the "bogus" matchings have little effect, as will be described shortly.

Let $\mathcal{M}$ denote the set of perfect matchings of $K_{n,n}$ and for $u \in V_1$ and $v \in V_2$ we let $\mathcal{M}(u, v)$ denote the set of near perfect matchings of $K_{n,n}$ that leave only u, v isolated. We are now ready to specify the desired stationary distribution of our Markov chain. This will be the distribution π over Ω defined by $\pi(M) \propto \Lambda(M)$, where

$$\Lambda(M) = \begin{cases} \lambda(M)w(u,v) & \text{if } M \in \mathcal{M}(u,v) \text{ for some } u, v; \\ \lambda(M) & \text{if } M \in \mathcal{M}, \end{cases}$$

and $w : V_1 \times V_2 \to \mathbb{R}^+$ is the weight function for *holes* to be specified shortly.

To construct a Markov chain having π as its stationary distribution, we use the original chain of [Bro86, JS89] augmented with a Metropolis acceptance rule for the transitions. Thus transitions from a matching M are defined as follows:

1. Choose an edge $e = (u, v)$ uniformly at random.
2. (i) If $M \in \mathcal{M}$ and $e \in M$, let $M' = M \setminus \{e\} \in \mathcal{M}(u, v)$;
   (ii) if $M \in \mathcal{M}(u, v)$, let $M' = M \cup \{e\} \in \mathcal{M}$;
   (iii) if $M \in \mathcal{M}(u, z)$ where $z \neq v$ and $(y, v) \in M$, let $M' = M \cup \{e\} \setminus \{(y, v)\} \in \mathcal{M}(y, z)$;

(iv) if $M \in \mathcal{M}(y, v)$ where $y \neq u$ and $(u, z) \in M$, let $M' = M \cup \{e\} \setminus \{(u, z)\} \in \mathcal{M}(y, z)$.

3. With probability $\min\{1, \Lambda(M')/\Lambda(M)\}$ go to M'; otherwise, stay at M.

The Metropolis rule in the final step ensures that this Markov chain is reversible with $\pi(M) \propto \Lambda(M)$ as its stationary distribution. Finally, to make the chain lazy we add a self-loop probability of 1/2 to every state; i.e., on every step, with probability 1/2 we make a transition as above and otherwise do nothing.

Next we need to specify the weight function w. Ideally we would like to take $w = w^*$, where

$$w^*(u, v) = \frac{\lambda(\mathcal{M})}{\lambda(\mathcal{M}(u, v))} \tag{3.11}$$

for each pair of holes u, v.

We will not be able to determine w^* exactly but will content ourselves with weights w satisfying

$$w^*(y, z)/2 \leq w(y, z) \leq 2w^*(y, z), \tag{3.12}$$

with very high probability.

The main technical result of this paper is the following theorem, which says that, provided the weight function w satisfies condition (3.12), the Markov chain is rapidly mixing. We present the theorem as it applies to an arbitrary bipartite graph, hence let $m = |E|$. Since we are working with $K_{n,n}$, for our purposes $m = n^2$.

Theorem 3.3.2 *Assuming the weight function w satisfies inequality (3.12) for all $(y, z) \in V_1 \times V_2$, then the mixing time of the Markov chain MC is bounded above by $\tau(\delta) = O(m^6 n^8 (n \log n + \log \delta^{-1}))$, provided the initial state is a perfect matching of maximum activity.*

Initially we have to take $\lambda(e) = 1$ for all $e \in V_1 \times V_2$ and $w(u, v) = n$ for all $u \in V_1, v \in V_2$. This a natural way of starting with parameters for which (3.12) holds. By a sequence of iterations to be described we are able to maintain (3.12) and at the same time reduce $\lambda(e)$ to at most $1/n!$ for all $e \notin E$.

At this point we see that if $\mathcal{M}'$ denotes the perfect matchings of G and $\mathcal{M}''$ denotes the perfect matchings of $K_{n,n}$ which are not in G then since $\Lambda(\mathcal{M}') = |\mathcal{M}'|$, $\Lambda(\mathcal{M}'') \leq 1$ and $\Lambda(\mathcal{M}(u, v)) \leq 2\Lambda(\mathcal{M})$ for all u, v, we see that

$$\frac{\Lambda(\mathcal{M}')}{\Lambda(\Omega)} \geq \frac{|\mathcal{M}'| - 1}{2(n^2 + 1)}. \tag{3.13}$$

It follows from this and Theorem 3.3.2 that we can in polynomial time generate a near uniform perfect matching of G. We repeat this process a number of times until we find an edge $e_1 = (u_1, v_1)$ whiich is in at least a fraction $1/(2n^2)$ of the matchings in $\mathcal{M}'$. We then estimate this proportion to within accuracy $\epsilon/(2n)$ with probability at least δ/n. Let us call our estimate ρ_1. We then apply the same strategy to $G - \{u_1, v_1\}$ and so on to obtain estimated proportions $\rho_1, \rho_2, \ldots, \rho_n$. Our final estimate of the number of perfect matchings in G is then $\rho_1^{-1}\rho_2^{-1}\cdots\rho_n^{-1}$.

Now let us see how to go about reducing $\lambda(e)$ for $e \notin E$. Assuming that (3.12) holds we get sufficient samples from our chain so that we can estimate all of the $w^*(y,z)$ to within a factor 4/3 say. Let these estimates be denoted by $w'(y,z)$. If now there is an edge $e \notin E$ such that $\lambda(e) > 1/n!$ then we replace $\lambda(e)$ by $3\lambda(e)/4$. The effect of this is to change any value of w^* by at most 4/3. If we replace our old w values by the corresponding w' values then (3.12) will still hold, since now $w'(y,z)/w^*(y,z) \in [9/16, 16/9]$ for all y, z and we repeat our attempts at reducing the λ's where necessary. Thus to prove Theorem 3.3.1, it is sufficient to prove Theorem 3.3.2.

3.3.1 Proof of Theorem 3.3.2

Theorem 3.3.3 *For an ergodic, reversible Markov chain with self-loop probabilities $P(y,y) \geq 1/2$ for all states y, and any initial state $x \in \Omega$,*

$$\tau_x(\delta) \leq \frac{2}{\Phi^2}\left(\ln \pi(x)^{-1} + \ln \delta^{-1}\right).$$

We bound the conductance by defining canonical paths $\gamma_{I,F}$ from all $I \in \Omega$ to all $F \in \mathcal{M}$. By upper bounding the maximum number of paths through any particular transition we will obtain a lower bound on the conductance. Using the fact that perfect matchings are likely under the stationary distribution, it will be sufficient to only consider a portion of particular canonical paths. Denote the set of all canonical paths by $\Gamma = \{\gamma_{I,F} : (I,F) \in \Omega \times \mathcal{M}\}$. Certain transitions on a canonical path will be deemed *chargeable*. For each transition t denote by

$$\mathrm{cp}(t) = \{(I,F) : \gamma_{I,F} \text{ contains } t \text{ as a chargeable transition}\}.$$

The canonical paths are defined by superimposing I and F. If $I \in \mathcal{M}$, then $I \oplus F$ consists of a collection of alternating cycles. We assume that the cycles are ordered in some canonical fashion; for example, having ordered the vertices, we may take as the first cycle the one that contains the least vertex in the order, as the second cycle the one that contains the least vertex amongst those remaining, and so on. Furthermore we assume that each cycle has a distinguished start vertex (e.g., the least in the order).

The canonical path $\gamma_{I,F}$ from $I \in \mathcal{M}$ to F is obtained by unwinding these cycles in the canonical order. A cycle $v_0 \sim v_1 \sim \ldots \sim v_{2k} = v_0$, where we assume w.l.o.g. that the

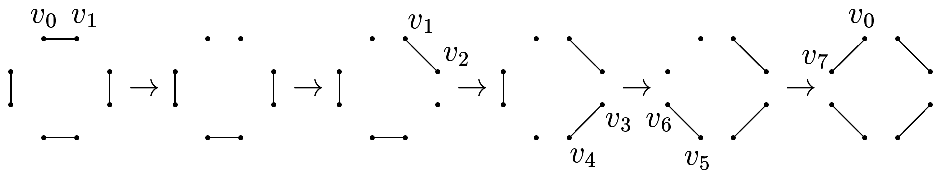

Figure 3.4: Unwinding a cycle with $k = 4$.

edge (v_0, v_1) belongs to I, is unwound by: (i) removing the edge (v_0, v_1), (ii) successively, for each $1 \le i \le k-1$, exchanging the edge (v_{2i}, v_{2i+1}) with (v_{2i-1}, v_{2i}), and (iii) adding the edge (v_{2k-1}, v_{2k}). (Refer to figure 3.4.) All transitions on the path $\gamma_{I,F}$ are deemed chargeable. A canonical path joining two perfect matchings, as just described, will be termed "type A."

If $I \in \mathcal{M}(y,z)$ for some $(y,z) \in V_1 \times V_2$, then $I \oplus F$ consists of a collection of alternating cycles together with a single alternating path from y to z. The canonical path $\gamma_{I,F}$ from I to F is obtained by unwinding the path and then unwinding the cycles in some canonical order. In this case, only the transitions involved in the unwinding of the path are deemed chargeable. The alternating path $y = v_0 \sim \ldots \sim v_{2k+1} = z$ is unwound by: (i) successively, for each $1 \le i \le k$, exchanging the edge (v_{2i-1}, v_{2i}) with (v_{2i-2}, v_{2i-1}), and (ii) adding the edge (v_{2k}, v_{2k+1}). A canonical path joining a near-perfect to a perfect matching will be termed "type B."

We define a notion of congestion of Γ that accounts only for the chargeable transitions:

$$\varrho(\Gamma) := \max_{t \in T} \left\{ \frac{1}{Q(t)} \sum_{(I,F) \in \mathrm{cp}(t)} \pi(I)\pi(F) \right\}. \tag{3.14}$$

Our main task will be to derive an upper bound on $\varrho(\Gamma)$, which we state in the next lemma. From this, it will be a straightforward matter to obtain a lower bound on the conductance Φ (see Lemma 3.3.2 below) and hence, via Theorem 3.3.3, a bound on the mixing time. In the following lemma recall that $m = |E|$, where for our purposes $m = n^2$.

Lemma 3.3.1 *Assuming the weight function w satisfies inequality (3.12) for all $(y,z) \in V_1 \times V_2$, then $\varrho(\Gamma) \le 16m$.*

In preparation for proving Lemma 3.3.1, we establish some combinatorial inequalities concerning weighted near-perfect matchings that will be used in the proof.

Lemma 3.3.2 *Let G be as above, and let $u, y \in V_1$, $v, z \in V_2$.*

1. *$\lambda(u,v)\lambda(\mathcal{M}(u,v)) \le \lambda(\mathcal{M})$, for all vertices u, v with $u \sim v$;*

2. $\lambda(u,v)\lambda(\mathcal{M}(u,z))\lambda(\mathcal{M}(y,v)) \le \lambda(\mathcal{M})\lambda(\mathcal{M}(y,z))$, *for all distinct vertices* u,v,y,z *with* $u \sim v$.

Proof The mapping from $\mathcal{M}(u,v)$ to $\mathcal{M}$ defined by $M \mapsto M \cup \{(u,v)\}$ is injective, and preserves activities modulo a factor $\lambda(u,v)$; this dispenses with (i). For (ii), suppose $M_{u,z} \in \mathcal{M}(u,z)$ and $M_{y,v} \in \mathcal{M}(y,v)$, and consider the superposition of $M_{u,z}$, $M_{y,v}$ and the single edge (u,v). Observe that $M_{u,z} \oplus M_{y,v} \oplus \{(u,v)\}$ decomposes into a collection of cycles together with an odd-length path O joining y and z.[3] Let $O = \{e_0, e_1, \dots, e_{2k}\}$ be an enumeration of the edges of this path, starting at y and working towards z. Denote by O_0 the $k+1$ even edges, and by O_1 the k odd edges. Finally define a mapping from $\mathcal{M}(u,z) \times \mathcal{M}(y,v)$ to $\mathcal{M} \times \mathcal{M}(y,z)$ by $(M_{u,z}, M_{y,v}) \mapsto (M, M_{y,z})$, where $M := M_{u,z} \cup O_0 \setminus O_1$ and $M_{y,z} := M_{y,v} \cup O_1 \setminus O_0$. Note that this mapping is injective, since we may uniquely recover $(M_{u,z}, M_{y,v})$ from $(M, M_{y,z})$. (To see this, observe that $M \oplus M_{y,z}$ decomposes into a number of cycles, together with a single odd-length path joining y and z. This path is exactly the path O considered in the forward map. There is only one way to apportion edges from $O \setminus \{(u,v)\}$ between $M_{u,z}$ and $M_{y,v}$.) Moreover, the mapping preserves activities modulo a factor $\lambda(u,v)$. □

Corollary 3.3.1 *Let G be as above, and let $u,y \in V_1$, $v,z \in V_2$. Then, provided in each case that the left hand side of the inequality is defined,*

1. $w^*(u,v) \ge \lambda(u,v)$, *for all vertices* u,v *with* $u \sim v$;
2. $w^*(u,z)w^*(y,v) \ge \lambda(u,v)w^*(y,z)$, *for all distinct vertices* u,v,y,z *with* $u \sim v$;
3. $w^*(u,z)w^*(y,v) \ge \lambda(u,v)\lambda(y,z)$, *for all distinct vertices* u,v,y,z *with* $u \sim v$ *and* $y \sim z$.

Proof Inequalities (i) and (ii) follow from the correspondingly labelled inequalities in Lemma 3.3.2, and the definition of w^*. Inequality (iii) is implied by (i) and (ii). □

Armed with Corollary 3.3.1, we can now turn to the proof of our main lemma.

Proof [Proof of Lemma 3.3.1] Note from the Metropolis rule that for any pair of states M, M' such that the probability of transition from M to M' is non-zero, we have $Q(M,M') = \min\{\pi(M), \pi(M')\}/2m$. We will show that for any transition $t = (M, M')$ and any pair of states $I, F \in \mathrm{cp}(t)$, we can define an encoding $\eta_t(I,F) \in \Omega$ such that $\eta_t : \mathrm{cp}(t) \to \Omega$ is an injection (i.e., (I,F) can be recovered uniquely from $\eta_t(I,F)$), and

$$\pi(I)\pi(F) \le 8\min\{\pi(M), \pi(M')\}\pi(\eta_t(I,F)) = 16m\,Q(t)\pi(\eta_t(I,F)). \tag{3.15}$$

[3] It is at this point that we rely crucially on the bipartiteness of G. If G is non-bipartite, we may end up with an even-length path and an odd-length cycle, and the proof cannot proceed.

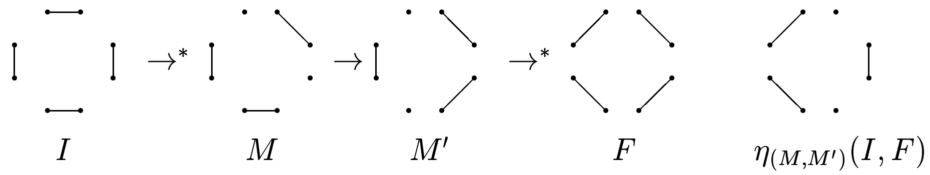

Figure 3.5: A canonical path through transition $M \to M'$ and its encoding.

Summing inequality (3.15) over $(I, F) \in \mathrm{cp}(t)$, we get

$$\frac{1}{Q(t)} \sum_{(I,F)\in\mathrm{cp}(t)} \pi(I)\pi(F) \leq 16m \sum_{(I,F)\in\mathrm{cp}(t)} \pi(\eta_t(I,F)) \leq 16m,$$

where we have used the fact that η_t is an injection. This immediately yields the claimed bound on $\varrho(\Gamma)$.

We now proceed to define the encoding η_t and show that it has the above properties. For a transition $t = (M, M')$ which is involved in stage (ii) of unwinding a cycle, the encoding is

$$\eta_t(I, F) = I \oplus F \oplus (M \cup M') \setminus \{(v_0, v_1)\}.$$

(Refer to figure 3.5, where just a single alternating cycle is viewed in isolation.) Otherwise, the encoding is

$$\eta_t(I, F) = I \oplus F \oplus (M \cup M').$$

It is not hard to check that $C = \eta_t(I, F)$ is always a matching in Ω (this is the reason that the edge (v_0, v_1) is removed in the first case above), and that η_t is an injection. To see this for the first case, note that $I \oplus F$ may be recovered from the identity $I \oplus F = (C \cup \{(v_0, v_1)\}) \oplus (M \cup M')$; the apportioning of edges between I and F can then be deduced from the canonical ordering of the cycles and the position of the transition t. The remaining edges, namely those in the intersection $I \cap F$, are determined by $I \cap F = M \cap M' \cap C$. The second case is similar, but without the need to reinstate the edge (v_0, v_1).

It therefore remains only to verify inequality (3.15) for our encoding η_t.

Consider first the case where $I \in \mathcal{M}$ and $t = (M, M')$ is the initial transition in the unwinding of an alternating cycle in a type A canonical path, where $M = M' \cup \{(v_0, v_1)\}$. Since $I, F, C, M \in \mathcal{M}$ and $M' \in \mathcal{M}(v_0, v_1)$, inequality (3.15) simplifies to

$$\lambda(I)\lambda(F) \leq 8 \min\{\lambda(M), \lambda(M')w(v_0, v_1)\}\lambda(C).$$

The inequality in this form can be seen to follow from the identity

$$\lambda(I)\lambda(F) = \lambda(M)\lambda(C) = \lambda(M')\lambda(v_0, v_1)\lambda(C),$$

using of inequality (i) of Corollary 3.3.1, and inequality (3.12). The situation is symmetric for the final transition in unwinding an alternating cycle.

Staying with the type A path, i.e., with the case $I \in \mathcal{M}$, suppose the transition $t = (M, M')$ is traversed in stage (ii) of unwinding an alternating cycle, i.e., exchanging edge (v_{2i}, v_{2i+1}) with (v_{2i-1}, v_{2i}). In this case we have $I, F \in \mathcal{M}$ while $M \in \mathcal{M}(v_0, v_{2i-1})$, $M' \in \mathcal{M}(v_0, v_{2i+1})$ and $C \in \mathcal{M}(v_{2i}, v_1)$. Since

$$\begin{aligned}\lambda(I)\lambda(F) &= \lambda(M)\lambda(C)\lambda(v_{2i}, v_{2i-1})\lambda(v_0, v_1)\\ &= \lambda(M')\lambda(C)\lambda(v_{2i}, v_{2i+1})\lambda(v_0, v_1),\end{aligned}$$

inequality (3.15) simplifies to

$$1 \le 8 \min\left\{\frac{w(v_0, v_{2i-1})}{\lambda(v_{2i}, v_{2i-1})}, \frac{w(v_0, v_{2i+1})}{\lambda(v_{2i}, v_{2i+1})}\right\} \frac{w(v_{2i}, v_1)}{\lambda(v_0, v_1)}.$$

This inequality can be verified by reference to Corollary 3.3.1: specifically, it follows from inequality (iii) in the general case $i \neq 1$, and by two applications of inequality (i) in the special case $i = 1$.

We now turn to the type B canonical paths. Suppose $I \in \mathcal{M}(y, z)$, and consider a transition $t = (M, M')$ from stage (i) of the unwinding of an alternating path, i.e., exchanging edge (v_{2i}, v_{2i-1}) with (v_{2i-2}, v_{2i-1}). Observe that $F \in \mathcal{M}, M \in \mathcal{M}(v_{2i-2}, z), M' \in \mathcal{M}(v_{2i}, z)$ and $C \in \mathcal{M}(y, v_{2i-1})$. Moreover, $\lambda(I)\lambda(F) = \lambda(M)\lambda(C)\lambda(v_{2i-2}, v_{2i-1}) = \lambda(M')\lambda(C)\lambda(v_{2i}, v_{2i-1})$. In inequality (3.15) we are left with

$$w(y, z) \le 8 \min\left\{\frac{w(v_{2i-2}, z)}{\lambda(v_{2i-2}, v_{2i-1})}, \frac{w(v_{2i}, z)}{\lambda(v_{2i}, v_{2i-1})}\right\} w(y, v_{2i-1}),$$

which holds by inequality (ii) of Corollary 3.3.1. Note that the factor $8 = 2^3$ is determined by this case, since we need to apply inequality (3.12) three times.

The final case is the last transition $t = (M, M')$ in unwinding an alternating path, where $M' = M \cup (z', z)$. Note that $I, C \in \mathcal{M}(y, z)$, $F, M' \in \mathcal{M}, M \in \mathcal{M}(z', z)$ and $\lambda(I)\lambda(F) = \lambda(M')\lambda(C) = \lambda(M)\lambda(z', z)\lambda(C)$. (Here we have written z' for v_{2k}.) Plugging these into inequality (3.15) leaves us with

$$1 \le 8 \min\left\{\frac{w(z', z)}{\lambda(z', z)}, 1\right\},$$

which follows from inequality (i) of Corollary 3.3.1.

We have thus shown that the encoding η_t satisfies inequality (3.15) in all cases. This completes the proof of the lemma. $\square$

Based on the upper bound on congestion, we can derive the following lower bound on the conductance.

Corollary 3.3.2 *Assuming the weight function w satisfies inequality (3.12) for all $(y,z) \in V_1 \times V_2$, then $\Phi \geq 1/100\varrho^3 n^4 \geq 1/10^6 m^3 n^4$.*

Proof Set $\alpha = 1/10\varrho n^2$. Let $\mathcal{S}, \overline{\mathcal{S}}$ be a partition of the state-space. (Note that we do not assume that $\pi(\mathcal{S}) \leq \pi(\overline{\mathcal{S}})$.) We distinguish two cases, depending on whether or not the perfect matchings $\mathcal{M}$ are fairly evenly distributed between $\mathcal{S}$ and $\overline{\mathcal{S}}$. If the distribution is fairly even, then we can show $\Phi(\mathcal{S})$ is large by considering type A canonical paths, and otherwise by using the type B paths.

Case I. $\pi(\mathcal{S}\cap\mathcal{M})/\pi(\mathcal{S}) \geq \alpha$ and $\pi(\overline{\mathcal{S}}\cap\mathcal{M})/\pi(\overline{\mathcal{S}}) \geq \alpha$. Just looking at canonical paths of type A we have a total flow of $\pi(\mathcal{S}\cap\mathcal{M})\pi(\overline{\mathcal{S}}\cap\mathcal{M}) \geq \alpha^2\pi(\mathcal{S})\pi(\overline{\mathcal{S}})$ across the cut. Thus $\varrho\, Q(\mathcal{S},\overline{\mathcal{S}}) \geq \alpha^2\pi(\mathcal{S})\pi(\overline{\mathcal{S}})$, and $\Phi(\mathcal{S}) \geq \alpha^2/\varrho = 1/100\varrho^3 n^4$.

Case II. Otherwise (say) $\pi(\mathcal{M}\cap\mathcal{S})/\pi(\mathcal{S}) < \alpha$. Note the following estimates:

$$\begin{aligned}
\pi(\mathcal{M}) &\geq \frac{1}{4n^2+1} \geq \frac{1}{5n^2};\\
\pi(\mathcal{S}\cap\mathcal{M}) &< \alpha\pi(\mathcal{S}) < \alpha;\\
\pi(\mathcal{S}\setminus\mathcal{M}) &= \pi(\mathcal{S}) - \pi(\mathcal{S}\cap\mathcal{M}) > (1-\alpha)\pi(\mathcal{S}).
\end{aligned}$$

Consider the cut $\mathcal{S}\setminus\mathcal{M} : \overline{\mathcal{S}}\cup\mathcal{M}$. The weight of canonical paths (all chargeable as they cross the cut) is $\pi(\mathcal{S}\setminus\mathcal{M})\pi(\mathcal{M}) \geq (1-\alpha)\pi(\mathcal{S})/5n^2 \geq \pi(\mathcal{S})/6n^2$. Hence $\varrho\, Q(\mathcal{S}\setminus\mathcal{M}, \overline{\mathcal{S}}\cup M) \geq \pi(\mathcal{S})/6n^2$. Noting $Q(\mathcal{S}\setminus\mathcal{M}, \mathcal{S}\cap M) \leq \pi(\mathcal{S}\cap M) \leq \alpha\pi(\mathcal{S})$ we have

$$\begin{aligned}
Q(\mathcal{S},\overline{\mathcal{S}}) &\geq Q(\mathcal{S}\setminus\mathcal{M},\overline{\mathcal{S}})\\
&= Q(\mathcal{S}\setminus\mathcal{M},\overline{\mathcal{S}}\cup M) - Q(\mathcal{S}\setminus\mathcal{M},\mathcal{S}\cap M)\\
&\geq (1/6\varrho n^2 - \alpha)\pi(\mathcal{S})\\
&\geq \pi(\mathcal{S})/15\varrho n^2\\
&\geq \pi(\mathcal{S})\pi(\overline{\mathcal{S}})/15\varrho n^2.
\end{aligned}$$

Rearranging, $\Phi(\mathcal{S}) = Q(\mathcal{S},\overline{\mathcal{S}})/\pi(\mathcal{S})\pi(\overline{\mathcal{S}}) \geq 1/15\varrho n^2$.

Clearly, it is Case I that dominates, giving the claimed bound on Φ. □

Theorem 3.3.2 of the previous section, now follows immediately.

Proof [Proof of Theorem 3.3.2] The condition on the starting state ensures $\log(\pi(X_0)^{-1}) = O(n\log n)$, where X_0 in the initial state. The lemma now follows from (2.1) and Theorem 2.2.1. □

initialize $\lambda(e) \leftarrow a_{\max}$ for all $e \in V_1 \times V_2$
initialize $w(u,v) \leftarrow na_{\max}$ for all $(u,v) \in V_1 \times V_2$
while $\exists e$ with $\lambda(e) > a(e)$ do
    take S samples from MC with parameters λ, w,
        each after a simulation of T steps
    use the sample to obtain estimates $w'(u,v)$ satisfying
        $3w^*(u,v)/4 \le w(u,v) \le 4w^*(u,v)/3$ with high probability $\forall u,v$
    set $\lambda(e) \leftarrow \max\{3\lambda(e)/4, a(e)\}$ and $w(u,v) \leftarrow w'(u,v)$ $\forall u,v$
output final weights $w(u,v)$

Figure 3.6: The algorithm for non-negative entries.

3.4 Arbitrary weights

Our algorithm easily extends to compute the permanent of a matrix A with non-negative entries. Let $a_{\max} = \max_{i,j} a(i,j)$ and $a_{\min} = \min_{i,j} a(i,j)$. Assuming $\mathrm{per}(A) > 0$, then it is clear that $\mathrm{per}(A) \ge (a_{\min})^n$. Rounding zero entries $a(i,j)$ to $(a_{\min})^n/n!$, the algorithm follows as described in figure 3.6.

The running time of this algorithm is polynomial in n and $\log(a_{\max}/a_{\min})$. For completeness, we provide a strongly polynomial time algorithm, i.e., one whose running time is polynomial in n and independent of $a_{\max}$ and $a_{\min}$, assuming arithmetic operations are treated as unit cost. The algorithm of Linial, Samorodnitsky and Wigderson [?] converts, in strongly polynomial time, the original matrix A into a nearly doubly stochastic matrix B such that $1 \ge \mathrm{per}(B) \ge \exp(-n - o(n))$ and $\mathrm{per}(B) = \alpha\,\mathrm{per}(A)$ where α is an easily computable function. Thus it suffices to consider the computation of $\mathrm{per}(B)$. In which case we can afford to round up any entries smaller than say n^{-2n} to n^{-2n}. The analysis for the 0,1-case now applies with the same running time.

3.5 Problems reducible to matchings

There are a number of problems reducible to counting and generating matchings. Here we discuss counting graphs with a fixed degree sequence, counting Hamilton cycles in *dense* graphs and counting Euler orientations.

Counting Graphs with a Fixed Degree Sequence

Let $\mathcal{G}(\mathbf{d})$ denote the set of all labelled graphs with vertex set $\{1, 2, \ldots, n\}$ and degree sequence $\mathbf{d} = (d_1, \ldots, d_n)$. It costs us very little in terms of complexity of discussion if we extend the discussion to allow the exclusion of a set of edges X. Thus let $\mathcal{G}(\mathbf{d}, X)$ denote the set of all graphs $G \in \mathcal{G}(\mathbf{d})$ for which the edge set of G is disjoint from X and so $\mathcal{G}(\mathbf{d}) = \mathcal{G}(\mathbf{d}, \emptyset)$. Our goal is to provide a fully polynomial almost uniform generator for $\mathcal{G}(\mathbf{d}, X)$, subject to appropriate conditions on $\mathbf{d}$ and X.

Define $\mathcal{G}'(\mathbf{d}, X)$ to be $\bigcup_{\mathbf{d}'} \mathcal{G}(\mathbf{d}', X)$, where the union ranges over vectors $\mathbf{d}' \in \mathbf{N}^n$ which satisfy $\mathbf{d}' \leq \mathbf{d}$ and $\sum_{i=1}^n |d_i - d_i'| \leq 2$. Call a class of degree sequences/excluded pairs *P-stable* if there exists a polynomial p such that $|\mathcal{G}'(\mathbf{d}, X)|/|\mathcal{G}(\mathbf{d}, X)| \leq p(n)$ for every sequence $\mathbf{d} = (d_1, \ldots, d_n)$ in the class. Informally, a degree sequence/excluded pairs $\mathbf{d}, X$ is P-stable if $|\mathcal{G}(\mathbf{d}, X)|$ does not change radically when $\mathbf{d}$ is slightly perturbed. Although the class of *all* graphical degree sequences (with $X = \emptyset$) is not P-stable, there are natural subclasses which are. We shall return to this issue in the next section.

Our aim is to construct a fully polynomial almost uniform generator for $\mathcal{G}(\mathbf{d}, X)$, which is valid for all sequences $\mathbf{d}$ within a specified P-stable class.

Theorem 3.5.1 *There is a good sampler and an FPRAS for $\mathcal{G}(\mathbf{d}, X)$ provided the pair $\mathbf{d}, X$ are drawn from some P-stable class.*

Proof For given degree sequence $\mathbf{d} = (d_1, \ldots, d_n)$ and excluded set X, let $\Gamma = \Gamma(\mathbf{d}, X)$ be the undirected graph with vertex set

$$V(\Gamma) = \{v_{ik} : 1 \leq i \leq n \text{ and } 1 \leq k \leq d_i\} \cup \{u_{ij} : 1 \leq i, j \leq n,\ i \neq j \text{ and } (i,j) \notin X\}$$

and edge set

$$\begin{aligned} E(\Gamma) = &\{(v_{ik}, u_{ij}) : 1 \leq i, j \leq n,\ 1 \leq k \leq d_i,\ i \neq j \text{ and } (i,j) \notin X\} \\ &\cup \{(u_{ij}, u_{ji}) : 1 \leq i, j \leq n,\ i \neq j \text{ and } (i,j) \notin X\}. \end{aligned}$$

The intention is to set up a correspondence between perfect matchings M in Γ and elements of $\mathcal{G}(\mathbf{d}, X)$. Informally, Γ contains an edge (u_{ij}, u_{ji}) corresponding to each potential edge (i,j) in a graph $G \in \mathcal{G}(\mathbf{d}, X)$; the *presence* of the edge (u_{ij}, u_{ji}) in M models the *absence* of the edge (i,j) in G. Additionally there are n *clusters* of vertices of the form $\{v_{ik} : 1 \leq k \leq d_i\}$ which, together with their incident edges, enforce the degree contraints at each vertex i in G.

Let ϕ be the function, from matchings in Γ to (undirected) graphs on vertex set $\{1, \ldots, n\}$, which maps the matching $M \subseteq E(\Gamma)$ to the graph with edge set

$$\{(i,j) : i \neq j \text{ and } (u_{ij}, u_{ji}) \notin M\}.$$

It is a straightforward task to verify that $\phi(\mathcal{M}) = \mathcal{G}(\mathbf{d}, X)$ and, moreover, that each graph in $\mathcal{G}(\mathbf{d}, X)$ is the image of precisely $\prod_{i=1}^n d_i!$ elements of $\mathcal{M}$. In particular, when $(u_{ij}, u_{ji}) \notin M$ there exist edges $(v_{is}, u_{ij}), (v_{jt}, u_{ji}) \in M$ and this enforces the degree constraints.

Thus, to generate elements of $\mathcal{G}(\mathbf{d}, X)$ almost uniformly, it is enough to generate perfect matchings in $\Gamma(\mathbf{d})$ almost uniformly. By Theorem 3.2.1, this will be possible provided

$|\mathcal{N}(\Gamma(\mathbf{d}))|/|\mathcal{M}(\Gamma(\mathbf{d}))| \le q(m)$, where $m = |V(\Gamma)|$ and q is some fixed polynomial. (The polynomial q will depend on the polynomial p in the definition of P-stable.)

Call a matching M of Γ *normalised* iff either (i) M is a perfect matching, or (ii) M is a near-perfect matching whose unmatched vertices are both cluster vertices. Denote the set of all normalised matchings of Γ by $\mathcal{N}'(\Gamma)$. Let $M \in \mathcal{N}(\Gamma)$ be a matching in which the vertex u_{ij} is unmatched. By adding the edge (u_{ij}, u_{ji}) to M, and removing the edge from M which was previously incident at u_{ji}, we succeed in moving one unmatched vertex into the set of cluster vertices. Two such operations are sufficient to normalise any near-perfect matching. (If vertices u_{ij} and u_{ji} are both unmatched, then M can be normalised by adding the single edge (u_{ij}, u_{ji}).)

The normalising operation maps at most n^2 distinct matchings onto a single normalised matching; hence $|\mathcal{N}(\Gamma)| \le n^2|\mathcal{N}'(\Gamma)|$. It is straightforward to check that $\phi(\mathcal{N}'(\Gamma)) = \mathcal{G}'(\mathbf{d}, X)$ and that each element of $\mathcal{G}'(\mathbf{d}, X)$ is the image of at most $\prod_{i=1}^n d_i!$ elements of $\mathcal{N}'(\Gamma)$. Putting these facts together we have

$$\frac{|\mathcal{N}(\Gamma)|}{|\mathcal{M}(\Gamma)|} \le \frac{n^2|\mathcal{N}'(\Gamma)|}{|\mathcal{M}(\Gamma)|} \le \frac{n^2|\mathcal{G}'(\mathbf{d}, X)|}{|\mathcal{G}(\mathbf{d}, X)|} \le n^2 p(n).$$

The proof is completed by appealing to Theorem 3.2.1; the degree of the polynomial q in the statement of that theorem can be taken as $\lceil \frac{1}{2} \deg p \rceil + 1$. □

The proof of Theorem 3.5.1 hinged on the fact that a polynomial bound on the ratio $|\mathcal{G}'(\mathbf{d}, X)|/|\mathcal{G}(\mathbf{d}, X)|$ implies a polynomial bound on the ratio $|\mathcal{N}(\Gamma(\mathbf{d}))|/|\mathcal{M}(\Gamma(\mathbf{d}))|$; in fact the reverse implication also holds. P-stability is therefore not merely a sufficient, but also a necessary condition for our reduction to be applicable.

In the next section, we will develop a simple numerical condition on degree sequences/excluded pairs which is sufficient to guarantee P-stability.

A criterion for P-stability

For a graphical degree sequence $\mathbf{d} = (d_1, \dots, d_n)$, define $d_{\max} = \max_i d_i$ and $e(\mathbf{d}) = \frac{1}{2}\sum_{i=1}^n d_i$. Note that $e(\mathbf{d})$ is integral. Next let $x_{\max}$ denote the maximum degree in the graph induced by X. We now derive a very useful sufficient condition for $\mathbf{d}$ to be P-stable in terms of the quantities $d_{\max}$, $e(\mathbf{d})$ and $x_{\max}$.

Theorem 3.5.2 *The class of all pairs $\mathbf{d}, X$ which satisfy $e(\mathbf{d}) > d_{\max}(d_{\max} + x_{\max} - 1)$ is P-stable.*

Proof Let $\mathbf{d}, X$ satisfy the above condition. We will show how to associate with each graph $G = (v, E) \in \mathcal{G}'(\mathbf{d}, X)$ a graph $\bar{G} \in \mathcal{G}(\mathbf{d}, X)$ which is "close to" G, in the sense that G can be transformed into $\bar{G}$ via a simple edge exchange operation. The result will then follow from the observation that no graph in $\mathcal{G}(\mathbf{d}, X)$ can be close to too many graphs in $\mathcal{G}'(\mathbf{d}, X)$.

If $G \in \mathcal{G}(\mathbf{d}, X)$ we simply set $\bar{G} = G$, so assume that $G \in \mathcal{G}'(\mathbf{d}, X) - \mathcal{G}(\mathbf{d}, X)$. We describe the operation first in the case that G has degree sequence $\mathbf{d}'$ of the form

$$d'_i = \begin{cases} d_i - 1 & \text{for } i \in \{k, l\}; \\ d_i & \text{otherwise.} \end{cases}$$

If $(k, l) \notin E \cup X$ then we just add this edge to G to form $\bar{G}$. If on the other hand the edge $(k, l) \in E \cup X$ we look for a pair x, y of vertices such that $(x, y) \in E$

(i) x, y, k, l are all distinct;

(ii) (x, k) and (y, l) are not edges of G.

(iii) $(x, k), (y, l) \notin X$.

Then the graph $\bar{G}$ is formed by adding to G the edges (x, k) and (y, l) and deleting the edge (x, y). We claim that such a pair x, y can always be found. To see this, note that there are $2e(\mathbf{d}') = 2(e(\mathbf{d}) - 1)$ candidates for the ordered pair x, y among endpoints of edges of G, some of which are excluded by requirements (i) – (iii). Elementary counting reveals that the number of candidates excluded by (i) is at most $2(d'_k + d'_l) - 2 \leq 2(2d_{\max} - 3)$. Similarly, the number excluded by (ii) is at most $2(d_{\max} - 2)(d_{\max} - 1)$ and the number excluded by (iii) is at most $2x_{\max}d_{\max}$. It follows that a suitable pair x, y can be found provided that

$$2(e(\mathbf{d}) - 1) > 2(2d_{\max} - 3) + 2(d_{\max} - 2)(d_{\max} - 1) + 2x_{\max}d_{\max} = 2d_{\max}(d_{\max} + x_{\max} - 1) - 2,$$

which is equivalent to the condition on $\mathbf{d}$ stipulated in the theorem.

It remains to describe $\bar{G}$ when the degree sequence of G is

$$d'_i = \begin{cases} d_i - 2 & \text{for } i = k; \\ d_i & \text{otherwise.} \end{cases}$$

In this case, we seek an edge (x, y) of G for which (i) $x, y \neq k$, and (ii) $(x, k), (y, k) \notin E \cup X$. The graph $\bar{G}$ is then obtained from G by adding the edges (x, k) and (y, k) and deleting (x, y). Using similar reasoning to the above, the reader may easily verify that a suitable edge (x, y) always exists under the stated condition on $\mathbf{d}$.

Now for any graph $H \in \mathcal{G}(\mathbf{d}, X)$, define the set

$$\mathcal{K}(H) = \{G \in \mathcal{G}'(\mathbf{d}, X) : \bar{G} = H\}.$$

Note that the sets $\mathcal{K}(H)$ partition $\mathcal{G}'(\mathbf{d}, X)$. It is a straightforward task to verify that each element of $\mathcal{K}(H)$ can be coded by a unique tuple (x, y, k, l), and hence that $|\mathcal{K}(H)| \leq n^4$. We therefore conclude that $|\mathcal{G}'(\mathbf{d}, X)| \,/\, |\mathcal{G}(\mathbf{d}, X)| \leq n^4$, so this class of degree sequences is indeed P-stable. $\square$

For the remainder of this discussion we consider $X = \emptyset$ and talk about degree sequences only. Informally, Theorem 3.5.2 says that a degree sequence belongs to a P-stable class provided its maximum and average degrees do not differ by too much. Let us mention two important types of degree sequence which satisfy this condition:

(i) Let $\mathbf{d} = (k, k, \ldots, k)$, i.e., $\mathbf{d}$ is k-regular, with $k \leq n/2$. Then setting $e(\mathbf{d}) = nk/2$ and $d_{\max} = k$, we see at once that the hypothesis of Theorem 3.5.2 is satisfied.

(ii) Suppose all degrees in $\mathbf{d}$ lie in the range $[1, \sqrt{n/2}\,]$. Then setting $e(\mathbf{d}) \geq \frac{1}{2}(d_{\max} + n - 1)$ and $d_{\max} \leq \sqrt{n/2}$, the hypothesis of Theorem 3.5.2 is again seen to hold.

Note also that there is an obvious bijection between the sets $\mathcal{G}(\mathbf{d})$ and $\mathcal{G}(\overline{\mathbf{d}})$, where $\overline{\mathbf{d}}$ is the complement of $\mathbf{d}$, i.e., $\overline{d}_i = n - 1 - d_i$. Hence for the purposes of generation and counting $\mathbf{d}$ and $\overline{\mathbf{d}}$ are equivalent.

Remark 3.5.1 The class of *all* graphical degree sequences is not P-stable. To see this, consider the family of degree sequences on $2k$ vertices of the form

$$\mathbf{d}^{(k)} = (1, 2, \ldots, k-1, k, k, k+1, \ldots, 2k-1)$$

for $k = 3, 4, \ldots$. Observe that the selection of a graph with any given degree sequence $\mathbf{d} = (d_1, \ldots, d_n)$ can be viewed recursively as follows:

1. Select a set of d_n neighbours for vertex n.

2. Recursively select a graph with degree sequence $\mathbf{d}' = (d'_1, \ldots, d'_{n-1})$, where $\mathbf{d}'$ is obtained from $\mathbf{d}$ by deleting the final component and decrementing by 1 the components which correspond to the d_n chosen neighbours for vertex n.

Adopting this view, it is clear that there is a unique graph with degree sequence $\mathbf{d}^{(k)}$.

Now modify $\mathbf{d}^{(k)}$ by decrementing by 1 the final two components. Note that graphs on the modified degree sequence are members of $\mathcal{G}'(\mathbf{d}^{(k)})$. Applying the recursive selection procedure we therefore have

$$\begin{aligned}
|\mathcal{G}'(\mathbf{d}^{(k)})| &\geq |\mathcal{G}(1,2,\dots,k,k,k+1,\dots,2k-3,2k-3,2k-2)| \\
&\geq |\mathcal{G}(1,1,2,\dots,k-1,k-1,k,\dots,2k-4,2k-4)| \\
&\geq |\mathcal{G}(1,1,1,2,\dots,k-2,k-2,k-1,\dots,2k-5)| \\
&\geq 3 \times |\mathcal{G}(1,1,1,2,\dots,k-3,k-3,k-2,\dots,2k-7)| \\
&\geq 3^2 \times |\mathcal{G}(1,1,1,2,\dots,k-4,k-4,k-3,\dots,2k-9)| \\
&\vdots \\
&\geq 3^{k-3}|\mathcal{G}(1,1,1,1)| \\
&= 3^{k-2}.
\end{aligned}$$

(The factor of 3 arises at each stage from the freedom to choose one of three degree-one vertices to be adjacent to the vertex of largest degree.) The ratio $|\mathcal{G}'(\mathbf{d}^{(k)})|/|\mathcal{G}(\mathbf{d}^{(k)})|$ is exponential in k, and hence in $n = 2k$, the number of vertices.

Remark 3.5.2 When considering bipartite graphs we can dispense with worrying about P-stability. Suppose we consider bipartite graphs with $m+n$ vertices. Then we change the definition of Γ so that it is bipartite, allowing the generation of a (near) random perfect matching without qualification.

$$\begin{aligned}
V(\Gamma) =&\{v_{ik} : 1 \leq i \leq m \text{ and } 1 \leq k \leq d_i\} \cup \{w_{jk} : 1 \leq j \leq n \text{ and } 1 \leq k \leq d_j\} \\
&\cup \{u_{ij}, u'_{ij} : 1 \leq i \leq m,\ 1 \leq j \leq n, (i,j) \notin X\}
\end{aligned}$$

and edge set

$$\begin{aligned}
E(\Gamma) = &\{(v_{ik}, u_{ij}) : 1 \leq i, \leq m,\ 1 \leq j \leq n\, 1 \leq k \leq d_i,\ \text{ and } (i,j) \notin X\} \\
&\cup \{(w_{jk}, u'_{ij}) : 1 \leq i \leq m,\ 1 \leq j \leq n\, 1 \leq k \leq d_i,\ \text{ and } (i,j) \notin X\} \\
&\cup \{(u_{ij}, u'_{ij}) : 1 \leq i, j \leq n,\ \text{ and } (i,j) \notin X\}.
\end{aligned}$$

Digraphs

Consider a directed graph $\overrightarrow{G} = (\overrightarrow{V}, \overrightarrow{E})$, where the in-degree (out-degree, respectively) of a vertex $v \in \overrightarrow{V}$ is denoted by $d_-(v)$ $(d_+(v))$. A *0,1-flow* is defined as a subset of edges $\overrightarrow{E'} \subset \overrightarrow{E}$ such that in the resulting subgraph $(\overrightarrow{V}, \overrightarrow{E'})$, $d_-(v) = d_+(v)$ for all $v \in \overrightarrow{V}$. Counting the number 0,1 flows is reducible to computing the 0,1 permanent of an undirected bipartite graph $G = (V, E)$ as follows.

The graph $G = (V, E)$ consists of:

$$V = \left\{ \begin{array}{ll} h_{i,j}, m_{i,j}, t_{i,j} & \text{for all } \overrightarrow{v_i v_j} \in \overrightarrow{E}, \\ u_i^1, \ldots, u_i^{d_-(v_i)} & \text{for all } v_i \in \overrightarrow{V} \end{array} \right\},$$

$$E = \left\{ \begin{array}{ll} (h_{i,j}, m_{i,j}), (m_{i,j}, t_{i,j}) & \text{for all } \overrightarrow{v_i v_j} \in \overrightarrow{E}, \\ (u_i^k, h_{i,j}) & \text{for all } i, j, k \text{ where } u_i^k, h_{i,j} \in \overrightarrow{V}, \\ (u_i^k, t_{j,i}) & \text{for all } i, j, k \text{ where } u_i^k, t_{j,i} \in \overrightarrow{V} \end{array} \right\}$$

A $0,1$-flow $\overrightarrow{E'}$ is mapped to a perfect matching M in the following manner. For each $\overrightarrow{v_i v_j} \in \overrightarrow{E'}$ add the edge $(h_{i,j}, m_{i,j})$ to M, while for each $\overrightarrow{v_i v_j} \in \overrightarrow{E} \setminus \overrightarrow{E'}$ add the edge $(m_{i,j}, t_{i,j})$ to M. Now for $v_i \in \overrightarrow{V}$, observe that the set of vertices $\{h_{i,j}\}_j \cup \{t_{j',i}\}_{j'}$, consists of exactly $d_-(v_i)$ unmatched vertices. There are $d_-(v_i)!$ ways of pairing these unmatched vertices with the set of vertices $\{u_i^k\}_k$. Thus the flow E' corresponds to $\prod_{v \in \overrightarrow{V}} d_-(v)!$ perfect matchings of G, and it is clear that the mapping is a bijection. This implies the following corollary.

Corollary 3.5.1 *For an arbitrary directed graph $\overrightarrow{G}$, there exists an fpras for counting the number of 0,1 flows.*

Suppose the directed graph $\overrightarrow{G}$ has a fixed source s and sink t. After adding a simple gadget from t to s we can estimate the number of maximum $0,1$ flows from s to t by estimating the number of $0,1$ flows in the resulting graph.

Hamilton Cycles in Dense Graphs

Let $G = (V, E)$ be a graph, where $V = \{v_1, v_2, \ldots, v_n\}$. Denote the degree of vertex v_i by d_i, for $i = 1, 2, \ldots, n$. We will say that G is *dense* if $\min_i d_i \geq (\frac{1}{2} + \alpha)n$, where $0 < \alpha \leq \frac{1}{2}$ is a fixed constant. Under these circumstances it is known that G must contain a Hamilton cycle. Moreover, the proof of this fact is easily modified to give a simple polynomial-time algorithm for constructing such a Hamilton cycle. This algorithm, which uses edges whose existence is guaranteed by the pigeonhole principle to "patch together" disjoint cycles, provides the required easy decision procedure.

We consider here the natural but more difficult problems of counting the *number* of Hamilton paths and cycles in such graphs. These problems are in fact #P-complete, so exact counting is presumably intractable. More positively, our main results in Sections 3.5 and 3.5 establish the existence of *fpras's* for these counting problems when $\alpha > 0$. We may observe that if the degree condition is relaxed to $\min_i d_i \geq (\frac{1}{2} - \alpha_n)n$ with $\alpha_n = \Omega(n^{\kappa-1})$ for any fixed $\kappa > 0$, then the question of the existence of any Hamilton path or cycle becomes NP-Complete and so approximate counting is NP-hard. This is true even if we insist on G being k-connected for any $k = o(n)$. Start with an arbitrary graph G and add a clique C of size $m = n^{1/\kappa}$ and an independent set I of size $m - 1$ and then join every vertex in C to every other vertex, to produce a graph Γ. Then G has a Hamilton path if and only if Γ has a Hamilton cycle. Also Γ contains a Hamilton path if and only if G contains two vertex disjoint paths that cover all its vertices.

Thus our results establish quite precisely the difficulty of the counting problem except in the region where α is close to zero. Section 5 extends the positive results of the earlier sections to cover self-avoiding paths and cycles of all lengths.

The natural approach given previous successes in this area is to try to find a rapidly mixing Markov chain with state space the set of Hamilton cycles of a given dense graph, and possibly its Hamilton paths as well. Earlier attempts with this approach have proved fruitless. Somewhat surprisingly, the key lies in the fact that in dense graphs, Hamilton cycles form a substantial fraction of the set of 2-factors, a *2-factor* being defined as a set of vertex-disjoint cycles which together contain all vertices of G. This is not obvious a priori and the main technical difficulty in the approach lies in obtaining a good upper bound on the ratio of 2-factors to Hamilton cycles in a dense graph. A direct attack — relating the number of 2-factors with k cycles to the number with $k+1$ cycles — appears unworkable. Instead, we introduce a weight function on 2-factors that allows us to argue about the distribution of total weight as a function of the number of cycles. By a rather delicate analysis, we are able to show that the Hamilton cycles carry sufficient weight for our purpose. In summary we prove

Theorem 3.5.3 *If G is dense then there are fpras's for*

(a) *approximating its number of Hamilton cycles,*

(b) *approximating its number of Hamilton paths,*

(c) *approximating its number of cycles of all sizes,*

(d) *approximating its number of paths of all sizes.*

Outline Approach

Our approach to constructing an fpras for Hamilton cycles in a dense graph G is via a randomized reduction to sampling and estimating *2-factors* in G. Using the results of Section 3.5 we prove

Theorem 3.5.4 *There exist both a good sampler and an FPRAS for the set of 2-factors in a dense graph.*

Proof The set of 2-factors in a graph $G = (V, E)$ is equal to $\mathcal{G}(\mathbf{d}, X)$, where $\mathbf{d} = (2, 2, \ldots, 2)$, and $X = V^{(2)} - E$ is the complementary edge set to E. The result now follows from Theorems 3.5.1 and 3.5.2, since, for a dense G and n sufficiently large, $d_{\max} = 2$, $x_{\max} < \frac{1}{2}n - 1$, and $d_{\max}(d_{\max} + x_{\max} - 1) < n = e(\mathbf{d})$. □

Given Theorem 3.5.4, the reduction from Hamilton cycles to perfect matchings is easy to describe. We estimate first the number of 2-factors in G, and then the proportion of 2-factors which are Hamilton cycles. Both counting and sampling phases run in polynomial time, by Theorem 3.5.4, provided only that G is dense. For the sampling phase to also produce an accurate estimate of the number of Hamilton cycles, it is necessary that the ratio of 2-factors to Hamilton cycles in G not be too large, i.e. bounded by a polynomial in n. This will be established in Section 3.5.

Many 2-factors are Hamiltonian

Let n be a natural number and $\beta = 10/\alpha^2$. Let $k_0 = \lfloor \beta \ln n \rfloor$, and for $1 \le k \le n$, define $g(k) = n^\beta k!(\beta \ln n)^{-k}$, and

$$f(k) = \begin{cases} g(k), & \text{if } k \le k_0; \\ g(k_0), & \text{otherwise.} \end{cases}$$

Lemma 3.5.1 *Let f be the function defined above. Then*

1. *f is non-increasing and satisfies*
$$\min\{f(k-1), f(k-2)\} = f(k-1) \ge (\beta \ln n)k^{-1} f(k);$$
2. *$f(k) \ge 1$, for all k.*

Proof Observe that g is unimodal, and that k_0 is the value of k minimizing $g(k)$; it follows that f is non-increasing. When $k \le k_0$, we have $f(k-1) = g(k-1) = (\beta \ln n)k^{-1} g(k) = (\beta \ln n)k^{-1} f(k)$; otherwise, $f(k-1) = g(k_0) = f(k) \ge (\beta \ln n)k^{-1} f(k)$. In either case, the inequality in part 1 of the lemma holds.

Part 2 of the lemma follows from the chain of inequalities

$$\frac{1}{f(k)} \le \frac{1}{g(k_0)} \le \frac{(\beta \ln n)^{k_0}}{n^\beta k_0!} \le n^{-\beta} \sum_{k=0}^{\infty} \frac{(\beta \ln n)^k}{k!} = n^{-\beta} \exp(\beta \ln n) = 1.$$

□

Lemma 3.5.2 *Suppose α is constant greater than 0. Let $G = (V, E)$ be an undirected graph of order n and minimum degree $(\frac{1}{2} + \alpha)n$. Then the number of 2-factors in G exceeds the number of Hamilton cycles by at most a polynomial (in n) factor, the degree of the polynomial depending only on α.*

Proof For $1 \le k \le \lfloor n/3 \rfloor$, let Φ_k be the set of all 2-factors in G containing exactly k cycles, and let $\Phi = \cup_k \Phi_k$ be the set of all 2-factors. Define

$$\Psi = \{(F, F') : F \in \Phi_k, F' \in \Phi_{k'}, k' < k, \text{ and } F \oplus F' \cong C_6\},$$

where $\oplus$ denotes symmetric difference, and C_6 is the cycle on 6 vertices. Observe that (Φ, Ψ) is an acyclic directed graph; let us agree to call its component parts *nodes* and *arcs* to avoid confusion with the vertices and edges of G. Observe also that if $(F, F') \in \Psi$ is an arc, then F' can be obtained from F by deleting three edges and adding three others, and that this operation can decrease the number of cycles by at most two. Thus every arc $(F, F') \in \Psi$ is directed from a node F in some Φ_k to a node F' in Φ_{k-1} or Φ_{k-2}.

Our proof strategy is to define a positive weight function on the arc set Ψ such that the total weight of arcs leaving each node (2-factor) $F \in \Phi \setminus \Phi_1$ is at least one greater than the total weight of arcs entering F. This will imply that the total weight of arcs entering Φ_1 is an upper bound on the number of non-Hamilton 2-factors in G, and that the maximum total weight of arcs entering a single node in Φ_1 is an upper bound on the ratio $|\Phi \setminus \Phi_1|/|\Phi_1|$.

The weight function $w : \Psi \to \mathbf{R}^+$ we employ is defined as follows. For any arc (F, F') with $F' \in \Phi_k$: if the 2-factor F' is obtained from F by coalescing two cycles of lengths l_1 and l_2 into a single cycle of length $l_1 + l_2$, then $w(F, F') = (l_1^{-1} + l_2^{-1}) f(k)$; if F' results from coalescing three cycles of length l_1, l_2 and l_3 into a single one of length $l_1 + l_2 + l_3$, then $w(F, F') = (l_1^{-1} + l_2^{-1} + l_3^{-1}) f(k)$.

Let $F \in \Phi_k$ be a 2-factor with $k > 1$ cycles $\gamma_1, \gamma_2, \ldots, \gamma_k$, of lengths $n_1, n_2, \ldots, n_k$. We proceed to bound from below the total weight of arcs *leaving* F. For this purpose imagine that the cycles $\gamma_1, \gamma_2, \ldots, \gamma_k$ are oriented in some way, so that we can speak of each oriented edge (u, u') in some cycle γ_i as being "forward" or "backward". Since we are interested in obtaining a *lower* bound, it is enough to consider only arcs (F, F^+) from F of a certain kind: namely, those for which the 6-cycle $\gamma = F \oplus F^+$ is of the form $\gamma = (x, x', y, y', z, z')$, where $(x, x') \in F$ is a forward cycle edge, $(y, y') \in F$ is a forward edge in a cycle distinct from the first, and $(z, z') \in F$ is a backward cycle edge. The edge (z, z') may be in the same cycle as either (x, x') or (y, y'), or in a third cycle. Observe that (x', y), (y', z) and (z', x) must necessarily be edges of F^+. It is routine to check that any cycle $\gamma = (x, x', y, y', z, z')$ satisfying the above constraints does correspond to a valid arc from F. The fact that (z, z') is oriented in the opposite sense to (x, x') and (y, y') plays a crucial role in ensuring that the number of cycles decreases in the passage to F^+ when only two cycles involved.

First, we estimate the number of cycles γ for which (x, x') is contained in a particular cycle γ_i of F. We might say that γ is *rooted* at γ_i. Assume, for a moment, that the vertices x, x', y, y' have already been chosen. There are at least $(\frac{1}{2} + \alpha)n - 5$ ways to extend the path (x, x', y, y'), first to z and then to z', which are consistent with the rules given above; let Z' be the set of all vertices z' so reachable. Denote by $G(x)$ the set of vertices adjacent to x. The number of ways of completing the path (x, x', y, y') to a valid 6-cycle is at least

$$\begin{aligned}|G(x) \cap Z'| &\geq |G(x)| + |Z'| - n \\ &\geq (\tfrac{1}{2} + \alpha)n + [(\tfrac{1}{2} + \alpha)n - 5] - n \\ &= 2\alpha n - 5 \\ &\geq \alpha n,\end{aligned}$$

for n sufficiently large. A lower bound on the number of 6-cycles γ rooted at γ_i now follows easily: there are n_i choices for (x, x'); then at least $(\frac{1}{2} + \alpha)n - n_i$ choices for (y, y'); and finally — as we have just argued — at least αn ways to complete the cycle. Thus the total number of 6-cycles rooted at γ_i is at least $\alpha n n_i[(\frac{1}{2} + \alpha)n - n_i]$.

We are now poised to bound the total weight of arcs leaving F. Each arc (F, F^+) defined by a cycle γ rooted at γ_i has weight at least $n_i^{-1} \min\{f(k-1), f(k-2)\}$, which, by Lemma 3.5.1, is bounded below by $(\beta \ln n)(kn_i)^{-1} f(k)$. Thus the total weight of arcs leaving F is bounded as follows:

$$\sum_{F^+:(F,F^+)\in\Psi} w(F, F^+) \geq \sum_{i=1}^{k} \alpha n n_i[(\tfrac{1}{2} + \alpha)n - n_i] \frac{(\beta \ln n) f(k)}{kn_i} \tag{3.16}$$

$$\begin{aligned}&= \alpha n^2[(\tfrac{1}{2} + \alpha)k - 1] \frac{(\beta \ln n) f(k)}{k} \\ &\geq \alpha^2 \beta\, f(k) n^2 \ln n \\ &\geq 10 f(k) n^2 \ln n,\end{aligned} \tag{3.17}$$

where we have used the fact that $k \geq 2$. Note that the presence of a unique backward edge, namely (z, z'), ensures that each cycle γ has a distinguishable root, and hence that the arcs (F, F^+) were not overcounted in summation (3.16).

We now turn to the corresponding *upper* bound on the total weight of arcs $(F^-, F) \in \Psi$ *entering* F. It is straightforward to verify that the cycle $\gamma = (x, x', y, y', z, z') = F^- \oplus F$ must contain three edges — (x, x'), (y, y') and (z, z') — from a single cycle γ_i of F, the remaining edges coming from F^-. The labeling of vertices in γ can be made canonical in the following way: assume an ordering on vertices in V, and assign label x to the smallest vertex. The condition $(x, x') \in F$ uniquely identifies vertex x', and the labeling of the other vertices in the cycle γ follows.

Removing the three edges (x, x'), (y, y') and (z, z') from γ_i leaves a triple of simple paths of lengths (say) $a - 1$, $b - 1$ and $c - 1$: these lengths correspond (respectively) to the segment containing x, the segment containing x', and the remaining segment. Going round the cycle γ_i, starting at x' and ending at x, the vertices x, x', y, y', z, z' may appear in one of eight possible sequences:

$$x', y', y, z', z, x;$$

$$\begin{array}{l} x', z, z', y, y', x; \\ x', z, z', y', y, x; \\ x', z', z, y, y', x; \\ x', y', y, z, z', x; \\ x', y, y', z', z, x; \\ x', z', z, y', y, x; \\ x', y, y', z, z', x. \end{array}$$

For a given triple of lengths (a, b, c), each of the above sequences corresponds to at most n_i possible choices for the edges (x, x'), (y, y') and (z, z'), yielding a maximum of $8n_i$ in total. To see this, observe that the edge (x, x') may be chosen in n_i ways (minimality of x fixes the orientation of the edge), and that the choice of (x, x') combined with the information provided by the sequence completely determines the triple of edges.

The eight sequences divide into five possible cases, as the first four sequences lead to equivalent outcomes (covered by case 1 below). Taken in order, the five cases are:

1. For at most $4n_i$ of the choices for the edges (x, x'), (y, y') and (z, z'), $\gamma_i \oplus \gamma$ is a single cycle;

2. for at most n_i choices, $\gamma_i \oplus \gamma$ is a pair of cycles of lengths a and $b + c$;

3. for at most n_i choices, $\gamma_i \oplus \gamma$ is a pair of cycles of lengths b and $a + c$;

4. for at most n_i choices, $\gamma_i \oplus \gamma$ is a pair of cycles of lengths c and $a + b$;

5. for at most n_i choices, $\gamma_i \oplus \gamma$ is a triple of cycles of lengths a, b and c.

The first case does not yield an arc (F^-, F), since the number of cycles does not decrease when passing from $F^- = F \oplus \gamma$ to F, but the other four cases do have to be reckoned with.

The total weight of arcs entering F can be bounded above as follows:

$$\begin{aligned}
\sum_{F^-:(F^-,F)\in\Psi} w(F^-,F) &\le \sum_{i=1}^{k} n_i f(k) \sum_{\substack{a,b,c\ge 1\\ a+b+c=n_i}} \left[\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\right)+\right.\\
&\qquad \left.\left(\frac{1}{a}+\frac{1}{b+c}\right)+\left(\frac{1}{b}+\frac{1}{a+c}\right)+\left(\frac{1}{c}+\frac{1}{a+b}\right)\right]\\
&= \sum_{i=1}^{k} n_i f(k) \sum_{\substack{a,b,c\ge 1\\ a+b+c=n_i}} \left[\frac{6}{a}+\frac{3}{b+c}\right]\\
&\le \sum_{i=1}^{k} n_i f(k) n \sum_{a=1}^{n_i-1} \left[\frac{6}{a}+\frac{3}{n_i-a}\right]\\
&\le 9 f(k) n^2 H_n \qquad (3.18)
\end{aligned}$$

where $H_n = \sum_{i=1}^{n} i^{-1} \le \ln n + 1$ is the nth harmonic number. Combining inequalities (3.17) and (3.18), we have

$$\begin{aligned}
\sum_{F^+:(F,F^+)\in\Psi} w(F,F^+) - \sum_{F^-:(F^-,F)\in\Psi} w(F^-,F) &\ge 10 f(k) n^2 \ln n - 9 f(k) n^2 H_n\\
&\ge f(k) n^2 (\ln n - 9)\\
&\ge n^2(\ln n - 9),
\end{aligned}$$

where the final inequality is by Lemma 3.5.1. Thus the total weight of arcs leaving F exceeds the total weight of arcs entering by at least 1, provided n is sufficiently large. The number of non-Hamilton 2-factors $|\Phi \setminus \Phi_1|$ is bounded above by the total weight of arcs entering Φ_1, which in turn is bounded — see inequality (3.18) — by $|\Phi_1| \times 9f(1)n^2 H_n = |\Phi_1| \times O(n^{2+\beta})$. This establishes the lemma. □

Counting the number of cycles of all sizes

We will first consider approximating the total number of cycles in graphs with minimum degree $(\frac{1}{2}+\alpha)n$.

We first note that if we add a loop to each vertex and extend the definition of 2-factor to include loops as cycles of length one, then the argument of Section 3.5 may be extended to this case (note that we still forbid cycles of length two i.e. double edges). Thus there exists both a fully polynomial randomized approximation scheme and a fully polynomial almost uniform sampler for the set of *partial* 2-factors in a dense graph. Let a partial 2-factor be *cyclic* if it consists of a single cycle of length at least three and a collection of loops. Clearly the number of cyclic partial 2-factors is the same as the number of cycles.

The procedure for approximating the number of cycles of all sizes is as follows: we estimate first the number of partial 2-factors in G, and then the number of cyclic partial 2-factors by standard sampling methods as a proportion of the number of partial 2-factors. To produce an accurate estimate in polynomial time it is only necessary to show that the ratio of partial 2-factors to cyclic partial 2-factors is not too large. Let

$$\mathcal{F}_\ell = \{\text{partial 2-factors with } \ell \text{ loops}\}, \quad \text{and} \quad f_\ell = |\mathcal{F}_\ell|.$$

For a given $F \in \mathcal{F}_\ell$ let $L = \{\text{loops of } F\}$, which we will now identify with the corresponding set of vertices. For $v \in L$ let d_v denote the number of neighbours of v in L and $D = \sum_{v \in L} d_v$.

If $v \in L$ then there are at least $2\alpha n - 2d_v$ ways of adding v to a cycle C of F by deleting an edge (a, b) of C and adding edges $(a, v), (v, b)$. Indeed we go round each cycle C of F; if the successor b of a vertex a neighbouring v is also a neighbour of v, then it forms an (a, b, v) triangle. The number of such triangles is at least $2\alpha n - 2d_v$.

So in total there are at least

$$\sum_{v \in L}(2\alpha n - 2d_v) = 2\ell\alpha n - 2D \tag{3.19}$$

$$\geq 2\ell(\alpha n - (\ell - 1)) \tag{3.20}$$

such augmentations.

Suppose first that $\ell \leq \ell_1 = \lfloor \alpha n/2 \rfloor$. Then (3.20) gives at least $\ell\alpha n$ augmentations of $F \in \mathcal{F}_\ell$ to an $F' \in \mathcal{F}_{\ell-1}$. Each $F' \in \mathcal{F}_{\ell-1}$ arises in at most n ways and so

$$\frac{f_{\ell-1}}{f_\ell} \geq \alpha\ell.$$

Putting $\ell_0 = \lceil 2/\alpha \rceil$ we see that

$$f_{\ell_1} + f_{\ell_1 - 1} + \cdots + f_{\ell_0 + 1} \leq f_{\ell_0} \leq f_{\ell_0} + f_{\ell_0 - 1} + \cdots + f_0. \tag{3.21}$$

Suppose next that $\ell > \ell_1$. Note first that since a graph with r vertices and s edges contains at least $r - s + 1$ distinct cycles, we see that L contains at least

$$\frac{D}{2} - \ell + 1 \tag{3.22}$$

distinct cycles.

Adding a cycle C contained in L to F and removing $|C|$ loops gives us a 2-factor in $\mathcal{F}_{\ell'}$ where $\ell' < \ell$. From (3.19) and (3.22) we see that there are at least

$$\left(\frac{2\ell\alpha n - 2D}{4}\right)^+ + \left(\frac{D}{2} - \ell\right)^+ \geq \ell\left(\frac{\alpha n}{2} - 1\right) \tag{3.23}$$

$$\geq \frac{\ell\alpha n}{3} \tag{3.24}$$

augmentations of either sort from F. Each $F' \in \mathcal{F}_{<\ell}$ arises in at most $n+n$ ways (accounting for both ways of reducing L) and so

$$\begin{aligned} f_\ell &\leq \frac{6}{\alpha\ell}(f_{\ell-1}+f_{\ell-2}+\cdots+f_0) \\ &\leq \theta(f_{\ell-1}+f_{\ell-2}+\cdots+f_0), \end{aligned}$$

where $\theta = 12/(\alpha^2 n)$, assuming $\ell > \ell_1$.

Thus

$$\frac{f_\ell+f_{\ell-1}+\cdots+f_0}{f_{\ell-1}+f_{\ell-2}+\cdots+f_0} \leq 1+\theta$$

and so

$$f_\ell+f_{\ell-1}+\cdots+f_0 \leq (1+\theta)^{\ell-\ell_1}\Sigma_1, \tag{3.25}$$

where $\Sigma_1 = f_{\ell_1}+f_{\ell_1-1}+\cdots+f_0$. We weaken (3.25) to

$$\begin{aligned} f_{\ell_1+k} &\leq (1+\theta)^k\Sigma_1 \\ &\leq e^{12\alpha^{-2}}\Sigma_1. \end{aligned} \tag{3.26}$$

It follows from (3.21) and (3.26) that

$$\frac{f_0+f_1+\cdots+f_n}{f_0+f_1+\cdots+f_{\ell_0}} \leq 2+2ne^{12\alpha^{-2}}. \tag{3.27}$$

Now take an $F \in \mathcal{F}_\ell$ where $\ell \leq \ell_0$ and fix its set of loops L. The number of partial 2-factors with this same L is at most a polynomial factor, $p(n)$ say, of the number of cycles of size $n-\ell$ through $V \setminus L$, by the results of Section 3. (It is clear that because ℓ is small here, the required degree conditions are satisfied.) Thus, by (3.27), the ratio of partial 2-factors to cyclic partial 2-factors is $O(np(n))$ and we have proved the existence of an fpras for the number of cycles.

Paths and Hamilton Paths

We obtain an fpras for counting the number of Hamilton paths in the following way. We add a vertex v_0 and join it by an edge to every vertex of G. Call this new graph G^*. The number of Hamilton cycles in G^* is equal to the number of Hamilton paths in G. Since G^* is dense we can approximate the latter quantity by approximating the former.

Similarly, to estimate the number of paths of all lengths, we compute an estimate c^* for the number of cycles in G^* and an estimate ρ^* for the proportion ρ of cycles which contain v_0. Since the number of cycles containing v_0 is the number of paths in G, this provides an estimate $\rho^* c^*$ for the number of paths. Also, this will give us an fpras provided ρ is not too small. But clearly $\rho \geq 3/4$ and we are done.

Eulerian Orientations

A graph $G = (V, E)$ is Eulerian if it is connected and all of its vertex degrees $d(v)$, $v \in V$ are even. An orientation σ of G defines a digraph $D_\sigma = (V, A_\sigma)$ where the unoriented edge $\{u, v\} \in E$ is replaced either by (u, v) or (v, u) in A_σ. Thus there are precisely $2^{|E|}$ distinct orientations of G.

Let $d_\sigma^+(v), d_\sigma^-(v)$ denote the outdegree, indegree of vertex v under orientation σ. σ is an Eulerian orientation if $d_\sigma^+(v) = d_\sigma^-(v)$ for all $v \in V$.

Theorem 3.5.5 *There is a good sampler and an FPRAS for the Euler orientations of an Eulerian graph G.*

Proof We reduce the problem of sampling/counting Euler orientations to that of sampling/counting perfect matchings in an associated bipartite graph G'. The graph G' has vertex bipartition $V' = V_1 \cup V_2$ where

$$V_1 = \bigcup_{v \in V} X_v \text{ and } X_v = \{x_{v,e} : \ v \in e \in E\} \text{ for } v \in V.$$

$$V_2 = \{w_e : \ e \in E\} \cup \bigcup_{v \in V} Y_v \text{ and } Y_v = \{y_{v,i} : \ 1 \leq i \leq d(v)/2\} \text{ for } v \in V.$$

The edge set E' of G' is defined by

$$E' = \{\{x_{u,e}, w_e\} : \ u \in e \in E\} \cup \bigcup_{v \in V} X_v \times Y_v.$$

Let $n' = 2m = |V_1| = |V_2|$ where $m = |E|$. Let now $\mathcal{M}'$ denote the set of perfect matchings of G'. Let $\mathcal{P}_0$ denote the set of Eulerian orientations of G.

Lemma 3.5.3 *$\mathcal{M}'$ can be partitioned into $\mathcal{M}'_\sigma : \sigma \in \mathcal{P}_0$ so that $|\mathcal{M}'_\sigma| = \prod_{v \in V} (d(v)/2)!$ for all $\sigma \in \mathcal{P}_0$.*

It follows immediately that sampling/counting for $\mathcal{P}_0$ can be reduced to sampling/counting for $\mathcal{M}'$.

Proof of Lemma 3.5.3 Given a perfect matching M of G' we let $X_v^M = \{x_{v,e} \in X_v : \ \{x_{v,e}, w_e\} \in M\}$. Then define an orientation σ as follows: If $e = \{u, v\} \in E$ then put (u, v) into A_σ if $x_{u,e} \in X_u^M$ and put (v, u) into A_σ if $x_{v,e} \in X_v^M$. Exactly one of these is true, in order that w_e is covered by M. The orientation is Eulerian because $d_\sigma^+(v) = |X_v^M| = d(v)/2$.

Furthermore, there are $\nu = \prod_{v \in V} (d(v)/2)!$ different matchings for each fixed collection $X_v^M : v \in V$, all giving the same orientation. The construction is reversible i.e. given $\sigma \in \mathcal{P}_0$ we let $X_v^M = \{x_{v,e} : \ e \text{ is oriented away from } v\}$ for $v \in V$. $\square$