COMBINATORICA
Bolyai Society – Springer-Verlag

COMBINATORICA 17 (3) (1997) 369–391

ON THE STOCHASTIC INDEPENDENCE PROPERTIES OF HARD-CORE DISTRIBUTIONS

JEFF KAHN* and P. MARK KAYLL†

Received June 4, 1996

A probability measure p on the set $\mathcal{M}$ of matchings in a graph (or, more generally 2-bounded hypergraph) Γ is *hard-core* if for some $\lambda:\Gamma\to[0,\infty)$, the probability $p(M)$ of $M\in\mathcal{M}$ is proportional to $\prod_{A\in M}\lambda(A)$. We show that such distributions enjoy substantial approximate stochastic independence properties. This is based on showing that, with M chosen according to the hard-core distribution p, $\mathrm{MP}(\Gamma)$ the matching polytope of Γ, and $\delta>0$, if the vector of *marginals*, $(\Pr(A\in M):A$ an edge of $\Gamma)$, is in $(1-\delta)\mathrm{MP}(\Gamma)$, then the weights $\lambda(A)$ are bounded by some $\Lambda(\delta)$. This eventually implies, for example, that under the same assumption, with δ fixed, $\frac{\Pr(A,B\in M)}{\Pr(A\in M)\Pr(B\in M)}\to 1$ as the distance between $A,B\in\Gamma$ tends to infinity.

Thought to be of independent interest, our results have already been applied in the resolutions of several questions involving asymptotic behaviour of graphs and hypergraphs (see [14, 16], [11]–[13]).

1. Introduction

This paper is concerned with "hard-core" probability distributions p on the set $\mathcal{M}$ of matchings in a multigraph or, more generally, in a 2-bounded hypergraph Γ, and in particular with showing that such distributions exhibit a significant amount of approximate stochastic independence. (See below for omitted notation and terminology.) This phenomenon, mainly as developed in the present work, has proved to be of considerable utility in several applications of the probabilistic method; see [13, 16, 14, 11, 12] for a roughly chronological sequence of such applications. Consequently, the authors believe hard-core distributions have the potential to play a major rôle in the further development of this ubiquitous method (see, *e.g.*, [1, 27]).

For p a probability distribution on the set $\mathcal{M}$ of matchings of Γ, the *marginals* of p are the numbers $p_A=\Pr(A\in M)$, for $A\in\Gamma$. We sometimes write f_p, or simply

Mathematics Subject Classification (1991): 05C70, 05C65, 60C05; 52B12, 82B20.

* Supported in part by NSF.

† This work forms part of the author's doctoral dissertation [16]; see also [17]. The author gratefully acknowledges NSERC for partial support in the form of a 1967 Science and Engineering Scholarship.

f, for the vector $(p_A : A \in \Gamma) \in \mathbf{R}^\Gamma$ of marginals of p. The set of such vectors is a familiar object, the *matching polytope* of Γ:

(1) $\mathrm{MP}(\Gamma) = \mathrm{conv}\{\mathbf{1}_M : M \in \mathcal{M}\} = \{f_p : p \text{ a probability distribution on } \mathcal{M}\}.$

A *hard-core* distribution (h.c.d.) is any $p=p_\lambda$ given, for some $\lambda:\Gamma\to[0,\infty)$, by

$$p(M) = \prod_{A\in M} \lambda(A) \Big/ \sum_{M'\in\mathcal{M}} \left(\prod_{A'\in M'} \lambda(A')\right)$$

(*i.e.*, $p(M)$ is proportional to the numerator in the expression above). As will become evident below (Theorem 3.1), such distributions are sufficiently plentiful to render them of considerable potential usefulness, *provided one can work with them.* Such a goal is supported by our main result, which quantifies the "approximate independence" to which we alluded above:

(1.1) Theorem. *For every $\delta\in(0,1)$, $\eta>0$ and $\ell\geq1$, there exist $\varepsilon>0$ and an integer D such that the following is true. Let Γ be a multigraph and $\mathcal{M}$ the set of matchings in Γ. Let p be a hard-core distribution on $\mathcal{M}$ with marginals $f_p\in(1-\delta)\mathrm{MP}(\Gamma)$, and let $\Gamma'=\{A\in\Gamma:p_A>\varepsilon\}$. If $F_1,\dots,F_\ell\in\Gamma$ are pairwise at distance at least D in Γ', then $\Pr(F_1,\dots,F_\ell\in M)=_\eta\prod_{i=1}^{\ell}p_{F_i}$.*

This rather technical form is for application in [14, 16]; for a more natural, though weaker, statement, take $\Gamma'=\Gamma$.

Of course, one sharpens the estimate by decreasing η, which must be accompanied by a suitable strengthening of the hypothesis. This is controlled by the parameters ε, D (chosen so that ε, D^{-1} are sufficiently small for δ, η, ℓ), since the condition "pairwise at distance at least D in Γ'" becomes increasingly difficult to satisfy as ε is decreased or D is increased.

Theorem 1.1 is based mainly on Lemma 4.1 and Corollary 4.2, which bound the edge weights $\lambda(A)$ in terms of δ (δ as in the statement of the theorem). Actually it might be more proper to say Lemma 4.1 is our "main" result, but we have given the name to Theorem 1.1 as it better embodies the properties of hard-core distributions we wish to bring to light. A primitive manifestation of these properties appears in Lemma 7.1, to which the reader might turn for an early and easily digested glimpse at approximate independence; indeed, this lemma forms a basis for the deeper Lemma 4.1 and may be of interest in its own right.

We pause to review terminology and fix notation.

Hypergraphs

Recall that a *hypergraph* $\mathcal{H}$ is a simply a collection, possibly with repeats, of nonempty subsets (*edges*) of some finite set $V=V(\mathcal{H})$ of *vertices*. A hypergraph is *k-uniform* (*k-bounded*) if each of its edges contains exactly (at most) k vertices.

The k-*edges* of $\mathcal{H}$ are the edges of cardinality k, and $\mathcal{H}_k$ denotes the set of these edges. A *multigraph* is a 2-uniform hypergraph; in particular, $\mathcal{H}_2$ is a multigraph for any $\mathcal{H}$. Hypergraphs in this paper are always 2-bounded (perhaps 2-uniform), and will usually be denoted by Γ or Γ'.

If $\mathcal{G} \subseteq \mathcal{H}$, then $\mathcal{G}$ is a *subhypergraph* of $\mathcal{H}$. For $X \subseteq V$, the subhypergraph $\mathcal{H}|_X$ of $\mathcal{H}$ *induced* by X is the hypergraph on X with edges $\mathcal{H}|_X = \{A \in \mathcal{H} : A \subseteq X\}$.

The *degree* in $\mathcal{H}$ of a vertex x is the number of edges containing x, and is denoted $d_{\mathcal{H}}(x)$.

We sometimes write $x \sim_\Gamma y$ to indicate that $\{x,y\} \in \Gamma$. In this case, we say that x is a *neighbour* of y in Γ, and vice versa. (Note we allow $x = y$ here.)

For $x, y \in V$, we use $\Delta(x,y) = \Delta_{\mathcal{H}}(x,y)$ for distance from x to y in $\mathcal{H}$, defined in the natural way. This extends routinely to distances between other types of objects in $\mathcal{H}$; for instance, if $\mathcal{F}, \mathcal{G} \subseteq \mathcal{H}$, then

$$\Delta(\mathcal{F}, \mathcal{G}) = \min\{\Delta(x,y) : x \in \cup_{A \in \mathcal{F}} A, y \in \cup_{B \in \mathcal{G}} B\}.$$

Matchings

A *matching* in a hypergraph is a set of pairwise disjoint edges; we use $\mathcal{M}(\mathcal{H})$ for the set of matchings of $\mathcal{H}$. The *matching polytope* of a (general) hypergraph $\mathcal{H}$ is defined analogously to (1):

$$\mathrm{MP}\,(\mathcal{H}) = \mathrm{conv}\,\{\mathbf{1}_M : M \in \mathcal{M}(\mathcal{H})\} \subseteq \mathbf{R}^{\mathcal{H}}$$

(with $\mathbf{1}_M$ the characteristic function of M). For $x \in V$ and $M \in \mathcal{M}(\mathcal{H})$, we write $x \prec M$ if x is contained in some element of M. In this case, we say x is *saturated* or *matched* by M.

Most of our graph and hypergraph notation follows [23] (see also [6]). A more general hypergraph reference is [3]; for omitted graph theory terminology, see, *e.g.*, [4].

Generalities

We write $\mathbf{R}^+$ for $[0,\infty)$, $[N]$ for $\{1,\dots,N\}$, $A \mathring{\cup} B$ for the disjoint union of A, B and $\binom{X}{N}$ for the collection of N-element subsets of a set X. For real-valued functions f, g (on a common domain), we write $f = \Omega(g)$ if f/g is bounded from below by a positive constant. For $\varepsilon > 0$, we write $f =_\varepsilon g$ for $(1+\varepsilon)^{-1} < f/g < 1+\varepsilon$.

Outline

The rest of the paper is organized as follows. Sections 2 and 3 provide background on Edmonds' theorem and hard-core distributions respectively. The brief Section 4 contains statements of the key Lemma 4.1 and two consequences, which provide the basis for the proof of Theorem 1.1 given in Section 5. We take an excursion in Section 6 to discuss h.c.d.'s on matchings in bipartite multigraphs. In Section 7, we prove a simple lemma which is the point of departure for most of our results on h.c.d.'s. Finally, the proofs of the lemmas of Section 4 are given in Section 8.

2. The matching polytope*

The celebrated "matching polytope theorem" of J. Edmonds ([5] or, *e.g.*, [23, 26]) may be regarded as characterizing those $f \in [0,1]^\Gamma$ which are marginal vectors of probability distributions on $\mathcal{M} = \mathcal{M}(\Gamma)$; indeed, this is the viewpoint in the present paper. The result gives the matching polytope in terms of linear inequalities:

(2.1) Theorem. *Let* Γ *be a 2-bounded hypergraph and* $f:\Gamma \to \mathbf{R}^+$. *Then* $f \in \mathrm{MP}(\Gamma)$ *if and only if*

$$\sum_{A \ni x} f(A) \le 1 \tag{2}$$

for every $x \in V$ *and*

$$\sum_{\substack{A \subseteq W \\ A \in \Gamma_2}} f(A) \le \left\lfloor \frac{|W|}{2} \right\rfloor \tag{3}$$

for every $W \subseteq V$.

(This is stated in [5] and elsewhere only for *2-uniform* Γ, but the present version is easily seen to be equivalent to the original; see Proposition 2.3 below.) We refer to (2), (3) as *Edmonds' constraints* on $f:\Gamma \to \mathbf{R}^+$. It is also convenient to have a name for the inequalities obtained from Edmonds' constraints by multiplying the right hand sides by $(1-\delta)$, where $\delta \in [0,1]$ is fixed. For such δ, the $\mathrm{Ed}(\delta)$ *constraints* on $f:\Gamma \to \mathbf{R}^+$ are

$$\sum_{A \ni x} f(A) \le 1 - \delta \tag{4}$$

for $x \in V$, and

$$\sum_{\substack{A \subseteq W \\ A \in \Gamma_2}} f(A) \le (1-\delta) \left\lfloor \frac{|W|}{2} \right\rfloor \tag{5}$$

for $W \subseteq V$. Of course, these characterize $(1-\delta)\mathrm{MP}(\Gamma)$ as (2), (3) characterize $\mathrm{MP}(\Gamma)$:

* This section is adapted from a similar discussion in [14].

(2.2) Corollary. *Let* $\delta \in [0,1]$, Γ *a 2-bounded hypergraph and* $f : \Gamma \to \mathbf{R}^+$. *Then* $f \in (1-\delta)\mathrm{MP}(\Gamma)$ *if and only if* f *satisfies* (4) *for every* $x \in V$ *and* (5) *for every* $W \subseteq V$. ∎

We will often refer to a probability distribution p on $\mathcal{M}$ as having the *property* $\mathrm{Ed}(\delta)$, meaning that its marginals $f = f_p$ satisfy the $\mathrm{Ed}(\delta)$ constraints (4) and (5) for all $x \in V$ and $W \subseteq V$.

Before moving to the next section, let us point out that — just as (3), (5) are the "interesting" conditions in Theorem 2.1 and Corollary 2.2 — the rôle of singletons in the matching polytope of *any* hypergraph is rather minor:

(2.3) Proposition. *Let* $\delta \in [0,1]$, $\mathcal{H}$ *a hypergraph and* $f : \mathcal{H} \to \mathbf{R}^+$. *Then* $f \in (1-\delta)\mathrm{MP}(\mathcal{H})$ *if and only if* $f|_{\mathcal{H}\setminus\mathcal{H}_1} \in (1-\delta)\mathrm{MP}(\mathcal{H}\setminus\mathcal{H}_1)$ *and* $\sum_{A \ni x} f(A) \le 1-\delta$ *for every* $x \in V(\mathcal{H})$. ∎

3. Hard-core distributions

The main purpose of this section is to record some easy facts about and present those earlier results on hard-core distributions that are important for our needs. We also give some general background and state the basic existence result for such distributions.

3.1. Background

The name "hard-core" is that given to such distributions in statistical physics (*e.g.* [2]), where the values $\lambda(F)$ are sometimes called *activities*. "Monomer-dimer system" and "exclusion model" are also used; see, *e.g.*, [9, 10], [19]. See also the rather extensive literature, *e.g.* [7, 8], on the "matchings polynomial".

Recently, h.c.d.'s have played a central rôle in several diverse contexts, including [22], [25], [20, 21], [14, 16], [11, 12], [13] and [15]. They are termed "normal populations" in [25], where they are applied to dynamical systems. In [20, 21] they are viewed as "canonical" convex representations of points in Euclidean space, and the probabilistic interpretation is only implicit. The references [11], [12], [14], [16], and to a lesser degree [13], [15] (all applications to different problems in asymptotic combinatorics), exploit the approximate independence of hard-core distributions developed in the present paper; see [18] for an overview of some of these applications.

Let $\mathcal{D} \subseteq [0,1]^{\mathcal{M}}$ denote the set of *all* probability distributions on $\mathcal{M}$. As observed at the beginning of Section 2, the question of existence of a $p \in \mathcal{D}$ with given marginals is answered by Theorem 2.1. The analogous question for h.c.d.'s is answered by the following result, proved independently in [25] and [20, 21]; a simple proof is sketched in [11].

(3.1) Theorem. *If $f \in \mathbf{R}^{\Gamma}$, $f \geq 0$, then there is a hard-core distribution p on $\mathcal{M}$ with marginals f if and only if the inequalities (2), (3) are strict for f, and in this case p is both the unique such h.c.d. and the (unique) entropy-maximizing distribution with marginals f.*

Actually the theorem holds for general (not necessarily 2-bounded) hypergraphs, if we replace the assumption on the inequalities (2), (3) by "$f \in (1-\delta)\mathrm{MP}(\mathcal{H})$ for some $\delta > 0$" (which by Corollary 2.2 gives an equivalent statement when Γ is 2-bounded).

Theorem 3.1 provides the opening for the use of h.c.d.'s in [14, 16], [11, 12], and though we do not need it in the present work (since our results always hypothesize a h.c.d.), we include it here as it shows why h.c.d.'s should be useful: if there is *some* distribution with a given set of marginals, then there is actually a *hard-core* distribution with essentially the same marginals; moreover, as suggested by entropy maximization, one may expect that among distributions with a given set of marginals, it is the h.c.d. (if any) that has the "best" independence properties.

3.2. Preliminaries

Suppose $p \in \mathcal{D}$, and let $M \in \mathcal{M}$ be chosen according to p. We use $\{\overline{x}\}$ to denote the event that x is not saturated by M, and natural variations of this notation are used without comment. (So, *e.g.*, for $x, y \in V$ and $M \in \mathcal{M}$ chosen randomly, we use $\{x, \overline{y}\}$ to denote the event $\{x \prec M, y \not\prec M\}$.) Braces are omitted whenever this does not lead to confusion. Though we usually think of p as a function in $\mathcal{D}$, we frequently use it to denote probabilities of other events in $2^{\mathcal{M}}$; for example, we write $p(x, \overline{y})$ for $\Pr(\{x, \overline{y}\})$.

Now let $p = p_\lambda$ be a h.c.d. on $\mathcal{M}$. Notice that for $A = \{x, y\} \in \Gamma$ (possibly with $x = y$), the natural correspondence between matchings leaving x, y unsaturated and matchings containing A gives

$$p_A = \lambda(A) p(\overline{x}, \overline{y}). \tag{6}$$

Thus, with $\lambda_{xy} = \sum_{A=\{x,y\}} \lambda(A)$, we have

$$p(\overline{x}) + \sum_{y} \lambda_{xy} p(\overline{x}, \overline{y}) = 1,$$

which, when divided by $p(\overline{x})$, gives the basic identity

$$p(\overline{x}) = \left[1 + \sum_{y} \lambda_{xy} p(\overline{y} | \overline{x})\right]^{-1}. \tag{7}$$

(We do not make much use of this here — it appears only in Section 6 — but see *e.g.* [11]–[13], [15], for a better appreciation of its significance.)

Observe that if $\Gamma' \subseteq \Gamma$ and $p = p_\lambda$, then p conditioned on $\{M \subseteq \Gamma'\}$ is just the h.c.d. on $\mathcal{M}(\Gamma')$ corresponding to $\lambda' := \lambda|_{\Gamma'}$ (sometimes termed the h.c.d. *induced* by Γ' or $\mathcal{M}(\Gamma')$). We use this observation (tacitly) in what follows, with $\Gamma' = \Gamma|_U$ for some $U \subseteq V$, in which case the h.c.d. induced by Γ' is simply $p\left(\,\cdot\,\middle|\bigwedge_{x\in V\setminus U}\overline{x}\right)$.

3.3. Mating

An important property of h.c.d.'s — see especially [25] — is that they are fixed points of the "mating map" $\varphi : \mathcal{D} \to \mathcal{D}$, defined based on the following method of generating new matchings from old. Suppose $p \in \mathcal{D}$, and let $M_1, M_2 \in \mathcal{M}$ be chosen independently according to p. Each component of the subhypergraph $M_1 \cup M_2$ of Γ has one of the following forms: a path or cycle in Γ_2 with successive edges alternating between M_1 and M_2; a path $\mathcal{P}$ as above, together with a singleton edge, from M_i of parity consistent with the 2-edges of $\mathcal{P}$, at one or both endpoints of $\mathcal{P}$ (*i.e.*, $\mathcal{P} \cup \{\{x\}\}$ or $\mathcal{P} \cup \{\{x\},\{y\}\}$, where $\mathcal{P}$ has endpoints x, y, and $\{x\} \in M_{3-i}$ if the 2-edge of $\mathcal{P}$ containing x lies in M_i, and similarly for y in the second case); a singleton $\{\{x\}\}$, where $\{x\} \in \Gamma$ lies in exactly one of M_1 or M_2; a singleton $\{F\}$, where $F \in \Gamma$ lies in both M_1 and M_2. Let M be a matching obtained from $M_1 \cup M_2$ by selecting, independently over all components, the M_1-members of a component with probability $\frac{1}{2}$, and otherwise selecting the M_2-members (so each member of $M_1 \cap M_2$ is necessarily contained in M). Denote the resulting probability distribution on $\mathcal{M}$ by $\varphi(p)$. The map $\varphi : \mathcal{D} \to \mathcal{D}$ so-defined is the *mating map* and we freely use the term *mating* when discussing the process of forming a (random) M from the randomly selected M_1, M_2.

Note that mating may be viewed as a method of generating a joint probability distribution on pairs of matchings, starting with some initial joint distribution on $\mathcal{M} \times \mathcal{M}$: given the (random) pair (M_1, M_2), mating produces a pair (M_1', M_2'), where M_1' is the M in the discussion defining φ, and M_2' is the complement of M_1' in the multiset $M_1 \cup M_2$. With this view of mating, we have the following observation from [25], where it was stated for 2-uniform Γ.

(3.2) Proposition. *Let M_1, M_2 be independent random matchings from a h.c.d. p, and let (M_1', M_2') be obtained from (M_1, M_2) by mating. Then (M_1, M_2) and (M_1', M_2') have the same joint distribution.*

Since $\mathcal{D}$ is a compact convex subset of the Euclidean space $\mathbf{R}^{\mathcal{M}}$, and $\varphi : \mathcal{D} \to \mathcal{D}$ is continuous, the Brouwer fixed-point theorem (see, *e.g.*, [24, Theorem 21.2]) implies that φ has at least one fixed point in $\mathcal{D}$. In fact, as the following restatement of Proposition 3.2 shows, φ has many fixed points.

(3.3) Corollary. *If $p \in \mathscr{D}$ is hard-core, then $\varphi(p)=p$.* ∎

For our purposes, this simple property of hard-core distributions is the crucial one: the properties we need from a hard-core p are established by regarding p as being generated from itself by mating (see in particular the sequence of claims in the proof of Lemma 8.1). A converse for Corollary 3.3 was also proved in [25].

4. Main lemma

Here we introduce the results on which Theorem 1.1 is based, postponing omitted proofs until Section 8. As usual, $p=p_\lambda$ denotes a h.c.d. on $\mathscr{M}(\Gamma)$.

As mentioned earlier, the next result is the heart the matter. It is also essential to [11] and [12].

(4.1) Lemma. *For any $\delta \in (0,1)$ there is a $\delta' = \delta'(\delta) \in (0,1)$ such that if p has property* Ed(δ), *then $p(\overline{x},\overline{y})>\delta'$ for all $x,y \in V$.*

In view of (6), this gives the weaker, but for our purposes essential (see Section 5),

(4.2) Corollary. *For any $\delta \in (0,1)$ there is a $\Lambda=\Lambda(\delta)$ such that if p has property* Ed(δ), *then $\lambda(A) \le \Lambda$ for all $A \in \Gamma$.* ∎

It should perhaps be emphasized that the functions δ', Λ delivered respectively by Lemma 4.1, Corollary 4.2 do not depend on Γ.

Remark. Corollary 4.2 raises an interesting general question. Recall that a collection $\mathscr{I}$ of subsets of a set X is an *ideal* of 2^X if $J \in \mathscr{I}$ whenever $J \subseteq I \in \mathscr{I}$ (so *e.g.* $\mathscr{M}(\Gamma)$ is an ideal of 2^Γ). The notions of hard-core distribution and matching polytope extend naturally to ideals (and more — see [20, 21]); namely, for $\lambda : X \to \mathbf{R}^+$, we may define $p=p_\lambda : \mathscr{I} \to [0,1]$ by

$$p(I) = \prod_{x \in I} \lambda(x) \Bigg/ \sum_{I' \in \mathscr{I}} \left(\prod_{x' \in I'} \lambda(x') \right),$$

while the analogue of the matching polytope is

$$K(\mathscr{I}) = \operatorname{conv}\{\mathbf{1}_I : I \in \mathscr{I}\}.$$

The (vague) question is then, what can be said about tradeoffs between λ and proximity of the vector f_p of marginals of p_λ to the boundary of $K(\mathscr{I})$? In particular, are there other interesting families of ideals $\mathscr{I}$ for which there is a $\Lambda=\Lambda(\delta)$ such that

$$f_p \in (1-\delta)K(\mathscr{I}) \Rightarrow \lambda(x) \le \Lambda \quad \text{for all } x \in X$$

(as in Corollary 4.2)?

The next result is also largely a consequence of Lemma 4.1.

(4.3) Lemma. *For any $\delta \in (0,1)$ there is a $\tilde{\delta} = \tilde{\delta}(\delta) \in (0,1)$ such that if p has property* $\mathrm{Ed}(\delta)$, *then $p(\cdot|\overline{x})$ has property* $\mathrm{Ed}(\tilde{\delta})$ *for all $x \in V$.*

5. Proof of main result

In this section, we prove the following superficially more general version of Theorem 1.1.

(5.1) Theorem. *For every $\delta \in (0,1)$, $\eta > 0$ and $\ell \geq 1$, there exist $\varepsilon > 0$ and an integer D such that the following is true. Let Γ be a 2-bounded hypergraph and $\mathcal{M} = \mathcal{M}(\Gamma)$. Let p be a h.c.d. on $\mathcal{M}$ with property* $\mathrm{Ed}(\delta)$, *and let $\Gamma' = \{A \in \Gamma : p_A > \varepsilon\}$. If $F_1, \ldots, F_\ell \in \Gamma$ are pairwise at distance at least D in Γ', then*

$$\Pr(F_1, \ldots, F_\ell \in M) =_\eta \prod_{i=1}^{\ell} p_{F_i}. \tag{8}$$

The basic idea for the proof of this is: if, as in Section 3.3, we view $M \in \mathcal{M}$, chosen according to p, as the result of mating the independently selected M_1, M_2 (each chosen according to p), then the only obstruction to *equality* in (8) is having a pair $F_i \neq F_j$ in the same component of $M_1 \cup M_2$. (To see this concretely, the reader might write down, *e.g.*, $\Pr(F_1, F_2 \in M)$ in terms of $\Pr(F_1 \in M_1, F_2 \in M_2)$, $\Pr(F_1 \in M_1, F_2 \notin M_2)$, *etc.* See [16] for a proof of Theorem 5.1 based explicitly on this idea. Here we use instead Theorem 5.4 from [2].) But, as shown in Lemma 5.2, the probability that F_i, F_j are in the same component of $M_1 \cup M_2$ is small.

For $F, G \in \Gamma$, write $\{F \sim_{1,2} G\}$ for the event that F, G lie in the same component of $M_1 \cup M_2$. If $M_1', M_2' \in \mathcal{M}$, use $\{F \sim_{1',2'} G\}$ similarly.

(5.2) Lemma. *For every $\varsigma > 0$ and $\delta \in (0,1)$, there exist $\varepsilon > 0$ and an integer D such that the following is true. If p, Γ, Γ' are as in the statement of Theorem 5.1, if M_1, M_2 are chosen independently according to p and if $F, G \in \Gamma$ satisfy $\Delta_{\Gamma'}(F,G) > D$, then $\Pr(F \sim_{1,2} G) < \varsigma p_F p_G$.*

Proof. Call the pair (M_1, M_2) *bad* in case the event $\{F \sim_{1,2} G\}$ holds for (M_1, M_2), and *good* otherwise. Let $\mathcal{B}$ denote the set of bad pairs and $\mathcal{A} = (\mathcal{M} \times \mathcal{M}) \setminus \mathcal{B}$ the set of good pairs. The idea of the proof is to show that

$$\Pr(\mathcal{B}) = \sum_{(M_1, M_2) \in \mathcal{B}} p(M_1) p(M_2) \tag{9}$$

is small compared to

$$\Pr(\mathcal{A}) = \sum_{(M_1', M_2') \in \mathcal{A}} p(M_1') p(M_2') \tag{10}$$

by defining a graph $\mathcal{G}$ with bipartition $(\mathcal{B},\mathcal{A})$ and (roughly) the following properties. First, certain terms in (9) (those we think of as large) correspond in $\mathcal{G}$ to many terms in (10), so each large term is overwhelmed by the number of corresponding terms. Second, each remaining (small) term in the first sum corresponds in $\mathcal{G}$ to a single, much larger term in the second. Third, each member of $\mathcal{A}$ has low degree (at most 6) in $\mathcal{G}$. This rough description, suitably quantified, will evidently give the desired

$$\Pr(\mathcal{B}) < \varsigma p_F p_G. \tag{11}$$

It is in implementing the description that we have our principal applications of Lemma 4.1; see Claim 5.3 and (13). (Lemma 4.1 appears again implicitly below in the inductive completion of the proof of Theorem 5.1, since this last step is based on Lemma 4.3.)

We define $\mathcal{G}$ by specifying the neighbours of each $(M_1,M_2) \in \mathcal{B}$. Let $\mathcal{B}_1 = \{(M_1,M_2) \in \mathcal{B} : \Delta_{M_1 \cup M_2}(F,G) > D\}$ and $\mathcal{B}_2 = \mathcal{B} \setminus \mathcal{B}_1$. Each $(M_1,M_2) \in \mathcal{B}_1$ is defined to be adjacent in $\mathcal{G}$ to all pairs $(M_1',M_2') \in \mathcal{A}$ obtained from (M_1,M_2) by deleting (i) a minimal set $\mathcal{C}$ of edges from $(M_1 \cup M_2) \setminus \{F,G\}$ for which F, G lie in different components of $(M_1 \cup M_2) \setminus \mathcal{C}$, and (ii) the edges F, G. Since the components of $M_1 \cup M_2$ are essentially paths and cycles (see the discussion in Section 3.3, which describes these components more carefully), no more than two edges of $(M_1 \cup M_2) \setminus \{F,G\}$ are deleted in obtaining any (M_1',M_2') from a given (M_1,M_2). Since $\Delta_{M_1 \cup M_2}(F,G) > D$, we have, for $(M_1,M_2) \in \mathcal{B}_1$,

$$d_{\mathcal{G}}((M_1,M_2)) > D. \tag{12}$$

(This can be increased to D^2 if F, G are on a cycle of $M_1 \cup M_2$.)

An *FG-path* is a (minimal) path $\mathcal{P}$ in Γ joining an element of F to an element of G. (The minimality ensures $F,G \notin \mathcal{P}$.) To define the neighbours of $(M_1,M_2) \in \mathcal{B}_2$, we need the following fact about FG-paths. We now specify $p = p_\lambda$.

(5.3) Claim. *Each FG-path $\mathcal{P}$ of length at most D contains an edge A with $\lambda(A) < \varepsilon/\delta'$, where $\delta' = \delta'(\delta)$ is the function appearing in Lemma 4.1.*

Proof. Since $\Delta_{\Gamma'}(F,G) > D$, there must be some $A = \{u,v\} \in \mathcal{P} \setminus \Gamma'$. The definition of Γ' gives $p_A \le \varepsilon$, so by (6) and our choice of δ' we have $\lambda(A) = p_A/p(\overline{u},\overline{v}) < \varepsilon/\delta'$. ∎

Each $(M_1,M_2) \in \mathcal{B}_2$ is then defined to be adjacent in $\mathcal{G}$ to a single $(M_1',M_2') \in \mathcal{A}$ obtained from (M_1,M_2) by deleting the following from $M_1 \cup M_2$: (i) an edge A with $\lambda(A) < \varepsilon/\delta'$ from an FG-path in $(M_1 \cup M_2) \setminus \{F,G\}$ of length at most D; (ii) the edges F, G; and, in case F, G are on a cycle of $M_1 \cup M_2$, (iii) an arbitrary edge from the only other FG-path in $(M_1 \cup M_2) \setminus \{F,G\}$.

Now compare the terms in the sums (9), (10). If $(M_1,M_2) \in \mathcal{B}$ and $(M_1,M_2) \sim_{\mathcal{G}} (M_1',M_2')$, then (M_1',M_2') was obtained from (M_1,M_2) by deleting, in addition to F and G, either a single edge A or a pair of edges A,B. The ratio $R :=$

$p(M_1)p(M_2)/p(M_1')p(M_2')$ is equal in the first case to $\lambda(A)\lambda(F)\lambda(G)$ and in the second to $\lambda(A)\lambda(B)\lambda(F)\lambda(G)$, so that, with $\Lambda=\Lambda(\delta)$ as in Corollary 4.2, we have $R\leq\Lambda^2\lambda(F)\lambda(G)$ (we may of course arrange that $\Lambda^2\geq\Lambda$). Thus,

$$p(M_1')p(M_2') \geq \frac{p(M_1)p(M_2)}{\Lambda^2\lambda(F)\lambda(G)} \tag{13}$$

(where we could replace Λ^2 by Λ if F, G are not on a cycle of $M_1 \cup M_2$). If, moreover, $(M_1,M_2)\in\mathcal{B}_2$, then we appeal to the construction of $\mathcal{G}$ to bound R from above by $\varepsilon\Lambda\lambda(F)\lambda(G)/\delta'$; that is,

$$p(M_1')p(M_2') \geq \frac{\delta'}{\varepsilon\Lambda\lambda(F)\lambda(G)}p(M_1)p(M_2). \tag{14}$$

The fact that $d_{\mathcal{G}}((M_1,M_2))=1$ for each $(M_1,M_2)\in\mathcal{B}_2$ and the inequalities (12)–(14) imply that

$$\sum_{\substack{(M_1',M_2')\sim_{\mathcal{G}}\\(M_1,M_2)}} p(M_1')p(M_2') \geq \tag{15}$$

$$\begin{cases} D\Lambda^{-2}(\lambda(F)\lambda(G))^{-1}p(M_1)p(M_2) & \text{if } (M_1,M_2)\in\mathcal{B}_1 \\ \varepsilon^{-1}\Lambda^{-1}\delta'(\lambda(F)\lambda(G))^{-1}p(M_1)p(M_2) & \text{if } (M_1,M_2)\in\mathcal{B}_2. \end{cases}$$

On the other hand, it is easy to see that $d_{\mathcal{G}}((M_1',M_2'))\leq 6$ for each $(M_1',M_2')\in\mathcal{A}$. Thus, setting $C=\max\{\Lambda^2/D,\varepsilon\Lambda/\delta'\}$ and applying (15), we have

$$\begin{aligned}\Pr(\mathcal{B}) &\leq C\lambda(F)\lambda(G)\sum_{(M_1,M_2)\in\mathcal{B}}\ \sum_{\substack{(M_1',M_2')\sim_{\mathcal{G}}\\(M_1,M_2)}} p(M_1')p(M_2') \\ &\leq 6C\lambda(F)\lambda(G)\Pr(\mathcal{A}).\end{aligned} \tag{16}$$

If we now choose D so large and ε so small that $6C<\varsigma\delta'^2$, then (6) and (16) (and again our choice of δ') give (11). ∎

Remaining details

We complete the proof of Theorem 5.1 by induction on $\ell\geq 1$. For $\ell=1$ there is nothing to prove. The induction will follow easily via Lemma 4.3 once we have the (main) case $\ell=2$. The latter is an immediate consequence of Lemma 5.2 and the following result of van den Berg and Steif [2], which expresses the covariance of the indicators $\mathbf{1}_{\{F\in M\}}$, $\mathbf{1}_{\{G\in M\}}$ in terms of probabilities in the product measure $p\times p$. In our language their result is

(5.4) Theorem. *For Γ, p as in Theorem 5.1, M chosen according to p, and M_1, M_2 chosen independently according to p,*

$$\Pr(F,G\in M)-p_Fp_G=\frac{1}{2}\Big[\Pr(F\sim_{1,2}G;\ F,G \text{ are in the same } M_i) - \Pr(F\sim_{1,2}G;\ F,G \text{ are in different } M_i\text{'s})\Big].$$

(This is really a special case; the full theorem ([2, Theorem 2.4]) applies to h.c.d.'s on independent sets of *vertices* in a graph.)

Now fix $\ell>2$, and assume we have Theorem 5.1 for each ℓ' with $1\le\ell'<\ell$. If $F_1=\{x,y\}\in\Gamma$ (possibly with $x=y$), then $p(\cdot|F_1\in M)=p(\cdot|\overline{x},\overline{y})$ (as distributions on $\mathcal{M}(\Gamma|_{V\setminus\{x,y\}})$). Thus, applying Lemma 4.3 (twice) shows that the (hard-core) distribution $p(\cdot|F_1\in M)$ enjoys the property $\mathrm{Ed}(\tilde\delta)$ for some $\tilde\delta=\tilde\delta(\delta)>0$. This and the induction hypothesis imply that for each $\eta'>0$, there are choices of ε, D to ensure that

$$\Pr(F_1,\ldots,F_\ell\in M)=_{\eta'}p_{F_1}\prod_{i=2}^{\ell}p(F_i\in M|F_1\in M)=_{\eta}\prod_{i=1}^{\ell}p_{F_i},$$

where $\eta>0$ can be made arbitrarily small through a suitable choice of η'. ∎

6. Bipartite multigraphs

Before turning to the proof of Lemma 4.1, we pause to consider the easier special case of *bipartite* Γ, for which we will focus on the bound $\Lambda=\Lambda(\delta)$ given in Corollary 4.2. Though the huge value produced by our proof of Lemma 4.1 (in Section 8) is adequate for present purposes (that is, we simply need *some* bound), it is clearly far from the truth, which seemingly should be that Λ is about δ^{-2}. We will show that, at least when Γ is bipartite, this is indeed the case; more precisely, we may take $\Lambda=1/(4\delta^2)$ (Proposition 6.3), and this is asymptotically best possible (Example 6.4).

We begin by verifying the plausible guess that deleting a vertex from one side of a bipartition of Γ never decreases the probability that a vertex on the opposite side remains unmatched:

(6.1) Lemma. *If Γ is a bipartite multigraph, $x,y\in V(\Gamma)$ are on opposite sides of a bipartition of Γ, and p is a h.c.d. on $\mathcal{M}(\Gamma)$, then*

$$p(\overline{x},\overline{y})\ge p(\overline{x})p(\overline{y}). \tag{17}$$

Proof. We use induction based on (7) to show that, more generally,

$$p(\overline{y}|\overline{x})\begin{cases}\le p(\overline{y}) & \text{if } x \text{ and } y,\ x\ne y, \text{ are on the same side}\\ \ge p(\overline{y}) & \text{if } x \text{ and } y \text{ are on opposite sides.}\end{cases} \tag{18}$$

Leaving basis cases for the reader to verify, we fix Γ and assume that the inequalities (18) hold for a h.c.d. on the matchings of any smaller Γ'.

Case 1. x and y, $x \neq y$, are on the same side of a bipartition of Γ.
Recall that the conditional distribution $p(\cdot|\overline{y})$ is again hard-core. Thus, the induction hypothesis shows that each $z \in V$ with $z \sim_\Gamma y$ (so that z is on the side opposite to x) satisfies $p(\overline{z}|\overline{x},\overline{y}) \geq p(\overline{z}|\overline{y})$, whence, using (7), we obtain

$$p(\overline{y}|\overline{x}) = \left[1 + \sum_{z \sim y} \lambda_{yz} p(\overline{z}|\overline{x},\overline{y})\right]^{-1} \leq \left[1 + \sum_{z \sim y} \lambda_{yz} p(\overline{z}|\overline{y})\right]^{-1} = p(\overline{y}).$$

Case 2. x and y are on opposite sides of a bipartition of Γ.
Now induction shows that each $z \in V$ with $x \neq z \sim_\Gamma y$ (so that z is on the same side as x) satisfies $p(\overline{z}|\overline{x},\overline{y}) \leq p(\overline{z}|\overline{y})$, yielding

$$p(\overline{y}|\overline{x}) = \left[1 + \sum_{x \neq z \sim y} \lambda_{yz} p(\overline{z}|\overline{x},\overline{y})\right]^{-1} \geq \left[1 + \sum_{z \sim y} \lambda_{yz} p(\overline{z}|\overline{y})\right]^{-1} = p(\overline{y}). \qquad \blacksquare$$

Lemma 6.1 (together with (4), (6)) gives

(6.2) Corollary. *If Γ is a bipartite multigraph and p is a h.c.d. on $\mathcal{M}(\Gamma)$ with property* Ed(δ)*, then each $A \in \Gamma$ satisfies $\lambda(A) \leq p_A/\delta^2 \leq 1/\delta^2$.* $\blacksquare$

A little extra effort yields the slightly stronger

(6.3) Proposition. *If Γ is a bipartite multigraph and p is a h.c.d. on $\mathcal{M}(\Gamma)$ with property* Ed(δ)*, then each $A \in \Gamma$ satisfies $\lambda(A) \leq p_A(1-p_A)/\delta^2 \leq 1/(4\delta^2)$.*

Proof. Let $A = \{x,y\}$. The Ed(δ) constraint (4) for x gives both $\delta \leq p(\overline{x}) = (1 - p_A)p(\overline{x}|A \notin M)$, and the analogous bound for y. Thus,

$$p(\overline{x}|A \notin M), p(\overline{y}|A \notin M) \geq \frac{\delta}{1-p_A}. \tag{19}$$

Since the conditional distribution $p(\cdot|A \notin M)$ is hard-core, Lemma 6.1 shows that

$$p(\overline{x},\overline{y}|A \notin M) \geq p(\overline{x}|A \notin M)p(\overline{y}|A \notin M). \tag{20}$$

Combining (19), (20) gives $p(\overline{x},\overline{y}) = (1-p_A)p(\overline{x},\overline{y}|A \notin M) \geq \delta^2/(1-p_A)$, and (6) then shows that $\lambda(A) \leq p_A(1-p_A)/\delta^2$. $\blacksquare$

Finally, as promised, we note that the bound of Proposition 6.3 is about right:

(6.4) Example. Let Γ be a cycle with m edges and $p = p_\lambda$ where $\lambda(A) = \lambda$ for each $A \in \Gamma$. Let δ_0 be the largest δ for which p satisfies the Ed(δ) constraints (4), (5) (namely $\delta_0 = p(\overline{x})$ — with x any vertex of Γ — if m is even, and $\delta_0 = p(\overline{x}) - (1-p(\overline{x}))/(2k)$ if $m = 2k+1$). Then a simple recursive calculation gives $\lim_{\lambda \to \infty} \lim_{m \to \infty} \delta_0 (4\lambda)^{1/2} = 1$. $\blacksquare$

7. A foundational lemma

We continue to let p denote a h.c.d. on $\mathcal{M}(\Gamma)$. The main result of this section, Lemma 7.1, says that the total perturbation in the quantities $p(y)$ caused by conditioning on $\{\overline{x}\}$ (or $\{x\}$) is small. As suggested in the introduction (Section 1), this is a first realization of the idea that hard-core distributions exhibit approximate independence properties. Though rather simple, this observation is the starting point for the more difficult results both of the present work and of [13].

If $w \in V$, then, as usual, $\Gamma - w$ denotes the hypergraph $\Gamma|_{V\setminus\{w\}}$. We also write $W - w$ for $W \setminus \{w\}$ when $W \subseteq V$. If $X \subseteq V$, then $\mu_\Gamma(X)$ denotes the expected number of vertices of X saturated by a matching M chosen according to p. The subscript Γ will often be replaced by Γ', an induced subhypergraph of Γ, in which case the expectation is understood to be taken with respect to the corresponding induced h.c.d. p'. Since the distribution will always be clear from context, explicit mention of p is omitted from the notation. We extend this notation to conditional expectations in the obvious way: for an event Ψ and $X \subseteq V$, the expected number of vertices of X saturated by M (chosen according to p), given Ψ, is denoted by $\mu_\Gamma(X|\Psi)$. Thus, for example, we have $\mu_\Gamma(W|\overline{w}) = \mu_{\Gamma-w}(W - w)$ for any $W \subseteq V$ and $w \in V$.

(7.1) Lemma. *For p a h.c.d. on $\mathcal{M}(\Gamma)$ and $W \subseteq V$:*
(a) *if $w \in W$, then $0 \le \mu_\Gamma(W) - \mu_{\Gamma-w}(W-w) = \mu_\Gamma(W) - \mu_\Gamma(W|\overline{w}) \le 2$;*
(b) *if $w \in V \setminus W$, then $|\mu_\Gamma(W) - \mu_{\Gamma-w}(W)| = |\mu_\Gamma(W) - \mu_\Gamma(W|\overline{w})| \le 1$.*

Proof. For the proof we write $p(uv)$ for $p(\{u,v\} \in M)$. We induct on $n = |V|$. The basis cases are trivial, so fix $n \ge 2$ and assume both (a) and (b) hold for Γ on fewer than n vertices. Each of the four inequalities for Γ with $|V(\Gamma)| = n$ depends on the inductive truth of both (a) and (b).
(a) Let $\{w_1, \ldots, w_a\}$ and $\{v_1, \ldots, v_b\}$ denote, respectively, the multisets of neighbours of w in $W - w$ and $V \setminus W$. We condition on whether w is saturated and further condition on which neighbour, if any (possibly including w itself), w is matched with:

$$\begin{aligned}\mu_\Gamma(W) - \mu_{\Gamma-w}(W-w) &= p(\overline{w})\mu_{\Gamma-w}(W-w) + p_{\{w\}}(1 + \mu_{\Gamma-w}(W-w)) \\ &\quad + \sum_{i=1}^{a} p(ww_i)\left[2 + \mu_{\Gamma-w-w_i}(W - w - w_i)\right] \\ &\quad + \sum_{j=1}^{b} p(wv_j)\left[1 + \mu_{\Gamma-w-v_j}(W-w)\right] - \mu_{\Gamma-w}(W-w) \\ &= \sum_{i=1}^{a} p(ww_i)\left[2 + \mu_{\Gamma-w-w_i}(W-w-w_i) - \mu_{\Gamma-w}(W-w)\right]\end{aligned}$$

$$(21) \qquad +\sum_{j=1}^{b} p(wv_j)\left[1+\mu_{\Gamma-w-v_j}(W-w)-\mu_{\Gamma-w}(W-w)\right]+p_{\{w\}}.$$

The induction hypothesis implies that

$$-2 \le \mu_{\Gamma-w-w_i}(W-w-w_i)-\mu_{\Gamma-w}(W-w) \le 0$$

for each $i \in [a]$, and

$$-1 \le \mu_{\Gamma-w-v_j}(W-w)-\mu_{\Gamma-w}(W-w) \le 1$$

for each $j \in [b]$. Thus, the expression in (21) is between $p_{\{w\}}$ and $2p(w \prec M, \{w\} \notin M)+p_{\{w\}}$, which is between 0 and 2.

(b) Defining $\{w_1,\ldots,w_a\}$ and $\{v_1,\ldots,v_b\}$ as in (a) (in this case each $v_j \neq w$) and conditioning as before, we have

$$\begin{aligned}\mu_\Gamma(W)-\mu_{\Gamma-w}(W) &= p(\overline{w})\mu_{\Gamma-w}(W)+p_{\{w\}}\mu_{\Gamma-w}(W)\\ &\quad+\sum_{i=1}^{a} p(ww_i)[1+\mu_{\Gamma-w-w_i}(W-w_i)]+\sum_{j=1}^{b} p(wv_j)\mu_{\Gamma-w-v_j}(W)-\mu_{\Gamma-w}(W)\\ &= \sum_{i=1}^{a} p(ww_i)[1+\mu_{\Gamma-w-w_i}(W-w_i)-\mu_{\Gamma-w}(W)]\\ &\quad+\sum_{j=1}^{b} p(wv_j)[\mu_{\Gamma-w-v_j}(W)-\mu_{\Gamma-w}(W)].\end{aligned}$$

The induction hypothesis now shows that the last expression is between -1 and 1. ∎

We are interested mainly in h.c.d.'s p with property Ed(δ) for a fixed $\delta \in (0,1)$. For such p and $x \in V$, the next result provides an upper bound for the number of $y \in V$ for which $p(\overline{x},\overline{y})$ (or $p(\overline{y}|\overline{x})$) is small. We again see approximate independence: conditioning on $\{\overline{x}\}$ can significantly affect only a few $p(\overline{y})$'s.

(7.2) Corollary. *If p is a h.c.d. with property* Ed(δ), *$x \in V$ and $W = \{y \in V : p(\overline{x},\overline{y}) < \delta^2/2\} \cup \{x\}$, then $|W| < 2/\delta$.*

Proof. Since p satisfies Ed(δ), we have $p(\overline{x}) \ge \delta$ and $p(\overline{y}|\overline{x}) < \delta/2$ for $y \in W \setminus \{x\}$. Thus,

$$\left(1-\frac{\delta}{2}\right)(|W|-1) < \mu_\Gamma(W|\overline{x}) \le \mu_\Gamma(W) \le (1-\delta)|W|$$

(using Lemma 7.1, Ed(δ) resp. for the second and third inequalities). The desired bound follows. ∎

8. Proofs of main lemmas

In this section, we (finally) give the proofs of Lemmas 4.1 and 4.3. As suggested earlier, this is perhaps the most interesting part of the paper. For Lemma 4.1, we need a uniform lower bound (*i.e.* depending only on δ) on the probabilities $p(\overline{x},\overline{y})$, for $x,y\in V$, when p is a h.c.d. with property Ed(δ). In practice we will assume (4) and, if some $p(\overline{x},\overline{y})$ is too small, show that some W violates (5).

Our strategy for identifying such a W is quite natural, at least in hindsight, though its implementation seems novel. If W is an odd subset of V for which (3) is close to exact, then the individual values $p(\overline{x})$, $x\in W$, may be large (on average they are at least $|W|^{-1}$ and in the interesting cases here W will not be very large); but the probabilities $p(\overline{x},\overline{y})$ with x,y distinct vertices of W should be small. This suggests fixing some x for which $p(\overline{x},\overline{y})$ is small for at least one y, and taking W to consist of x together with all such y (see (24)).

For technical reasons — we need a clear division between "small" and "not small" — our eventual definition of small (see β following (23)) is fairly drastic and leads to bounds which are surely far from the truth. In particular, it gives a rather enormous upper bound on the edge weights in Corollary 4.2, whereas, as suggested in Section 6, the correct bound is probably about δ^{-2}.

In proving Lemma 4.1, we find it convenient first to assume Γ has no singleton edges and then to extend to the general case.

(8.1) Lemma. *Lemma 4.1 holds in case Γ is a multigraph.*

Proof. Arguing by contradiction, we will demonstrate the existence of such a $\delta' \in (0,2^{-T^m}\delta^2)$, where $T=\Omega(\log\delta^{-1})$ and the integer $m=\Omega(\delta^{-2})$ are chosen to support the various inequalities arising in the proof.

If the result is false, then there exist $x,y_0\in V$ such that

$$p(\overline{x},\overline{y_0}) < 2^{-T^m}\delta^2. \tag{22}$$

(Of course, $x\neq y_0$ since p has property Ed(δ).) For $j=1,\ldots,m-1$, define intervals I_j by

$$I_j = \left[2^{-T^{(j+1)}}\delta^2, 2^{-T^j}\delta^2\right),$$

and for $y\in V$, define $Z_y\subseteq V$ by

$$Z_y = \left\{z : p(\overline{y},\overline{z}) < \delta^2/2\right\}\cup\{y\}.$$

Applying Corollary 7.2 to each Z_y gives $|Z_y| < 2\delta^{-1}$ for each $y\in V$. Thus, if $Z := Z_x$, then the number of pairs (y,z) with $y\in Z$ and $z\in Z_y$ is less than $4\delta^{-2}$, which our choice of m makes less than $m-1$. Since the number of intervals I_j is

$m-1$, some $j_0 \in [m-1]$ satisfies $p(\overline{y},\overline{z}) \notin I_{j_0}$ for all $y \in Z$ and $z \in Z_y$. If $z \in V \setminus Z_y$, then certainly $p(\overline{y},\overline{z}) \notin I_{j_0}$. Thus, for each $y \in Z$ and each $z \in V$, we have

$$p(\overline{y},\overline{z}) \notin I_{j_0}. \tag{23}$$

Write $I_{j_0} = [\beta, \alpha\beta)$, where $\beta := 2^{-T^{(j_0+1)}}\delta^2$ and $\alpha := 2^{T^{j_0}(T-1)}$. That I_{j_0} contains none of the probabilities $p(\overline{y},\overline{z})$ in (23) provides a useful device for sharpening both upper and lower bounds on these probabilities. For example, we can achieve $p(\overline{y},\overline{z}) < \beta$ by establishing the superficially weaker $p(\overline{y},\overline{z}) < \alpha\beta$; this observation is used at the end of the proof of Claim 1 below. Similarly, the bound $p(\overline{y},\overline{z}) \geq \alpha\beta$ follows once we obtain $p(\overline{y},\overline{z}) \geq \beta$; this is used in proving Claims 2 and 3 (see (30)).

Jumping a little ahead — indeed, the reader may wish to return to this remark after reaching (31) — we observe that the device in the preceding paragraph is essential, *e.g.*, for obtaining our estimate of the probability in Claim 2. Without the simultaneous sharpening of the relevant bounds, we would arrive at (31) equipped only to bound the probability there by (the useless) 2α, instead of the asserted $2\alpha^{-1}$. The proof of Claim 3 also calls (implicitly) for an application of these ideas.

In the remainder of the proof, it may help the reader to keep in mind the relative sizes of α, β and δ, which are, respectively, "huge", "tiny" and "small". To quantify this informal classification somewhat, we note that it will eventually suffice, *e.g.*, to arrange for each of β, α^{-1}, $\alpha^{-2}\beta^{-1}$ to be at most $\delta^3/24$. Since both α and β are functions of T — hence of the constant δ — they may be controlled through our choice of T. We will be more precise about our requirements of T when we reach (35), a glance at which may help to clarify the discussion in the present paragraph.

Our candidate for violating property $\mathrm{Ed}(\delta)$ is

$$W = \{y \in Z : p(\overline{x},\overline{y}) \leq \beta\} \cup \{x\}, \tag{24}$$

which, by (22), satisfies

$$|W| > 1 \tag{25}$$

((25) is needed in the proof of Claim 4). In showing, via the next four claims, that $\mathrm{Ed}(\delta)$ is indeed violated on W, we will several times apply Corollary 3.3, the fact that hard-core distributions are fixed points of the mating map.

Claim 1. If $u,v,w \in V$, with $v \neq w$, and either v or w is in Z, then the following implication holds: $p(\overline{u},\overline{v}), p(\overline{u},\overline{w}) \leq \beta \Rightarrow p(\overline{v},\overline{w}) \leq \beta$.

Proof. The implication is vacuous if either $u=v$ or $u=w$, so we assume that u, v and w are distinct. Generate a random $M \in \mathcal{M}$ (according to p) by mating the randomly (and independently) selected $M_1, M_2 \in \mathcal{M}$. The event $\Psi := \{\overline{u},\overline{v}\} \cup \{\overline{u},\overline{w}\}$ contains

$$\{u \not\sim M_1\} \cap \{v,w \not\sim M_2\} \cap E \cap F, \tag{26}$$

where E is the event that mating deletes the M_2-edge, if any, incident with u, and F the event that mating deletes the M_1-edge, if any, incident with v (unless u and v lie in the same component of $M_1 \cup M_2$, in which case u and w lie in different components — substitute w for v in this case). Thus,

$$p(\Psi) \geq p(\overline{u})p(\overline{v},\overline{w})/4. \tag{27}$$

From Ed(δ), we obtain a lower bound $\delta p(\overline{v},\overline{w})/4$ for the right side of (27), and the hypothesis $p(\overline{u},\overline{v}), p(\overline{u},\overline{w}) \leq \beta$ implies that $p(\Psi) \leq p(\overline{u},\overline{v}) + p(\overline{u},\overline{w}) \leq 2\beta$. Thus, we have $p(\overline{v},\overline{w}) \leq 8\beta/\delta$, and since our choice of T implies $8/\delta < \alpha$, it follows that $p(\overline{v},\overline{w}) < \alpha\beta$. By hypothesis, either v or w is in Z, so we may apply (23) to infer $p(\overline{v},\overline{w}) < \beta$, a stronger bound than we really need. ∎

In Claims 2–4, we write $\{z \prec_{\overline{W}} M\}$ for the event that the vertex $z \in W$ is matched with a vertex of $V \setminus W$ in the random matching M.

Claim 2. If $y, z \in W$ and $y \neq z$, then $p(y \not\prec M, z \prec_{\overline{W}} M) \leq 2\alpha^{-1}$.

Proof. The probability in question is

$$p(y \not\prec M, z \prec_{\overline{W}} M) = \sum_{w \notin W} p(y \not\prec M, \{w, z\} \in M), \tag{28}$$

and our bound arises from an argument similar to that in the preceding proof. Letting M_1, M_2 be chosen independently and at random according to p, we find

$$p(\overline{y},\overline{z}) \geq 2 \sum_{w \notin W} p(y \not\prec M_1, \{w, z\} \in M_1) p(w, z \not\prec M_2) \frac{1}{4}. \tag{29}$$

Since $y, z \in W \subseteq Z$, Claim 1 (with $u = x$) gives an upper bound β for the left side of (29). For $w \notin W$, Claim 1 implies that $p(\overline{w},\overline{z}) > \beta$; otherwise, the fact that $p(\overline{x},\overline{z}) \leq \beta$ would give $w \in W$. Thus, using (23), we see that each $w \in V \setminus W$ satisfies

$$p(w, z \not\prec M_2) \geq \alpha\beta. \tag{30}$$

Using these bounds in (29), we find

$$\beta \geq \frac{1}{2}\alpha\beta \sum_{w \notin W} p(y \not\prec M_1, \{w, z\} \in M_1). \tag{31}$$

The claim follows since the sums in (28), (31) are equal. ∎

Claim 3. If $y, z \in W$ and $y \neq z$, then $p(y \prec_{\overline{W}} M, z \prec_{\overline{W}} M) \leq 4\alpha^{-2}\beta^{-1}$.

Proof. If M_1, M_2 are chosen independently and at random according to p, and M is generated from (M_1, M_2) by mating, then

$$p(y \not\prec M, z \prec_{\overline{W}} M) \geq 2 \sum_{w \notin W} p(\{y, w\} \in M_1, z \prec_{\overline{W}} M_1) p(y, w \not\prec M_2) \frac{1}{4}. \tag{32}$$

Appealing to Claim 2, we observe that $2\alpha^{-1}$ is an upper bound for the left side of (32). Using (30), we see that each $w \in V \setminus W$ satisfies $p(y, w \not\sim M_2) \geq \alpha\beta$. Thus,

$$2\alpha^{-1} \geq \frac{1}{2}\alpha\beta \sum_{w \notin W} p(\{y,w\} \in M_1, z \prec_{\overline{W}} M_1) = \alpha\beta p(y \prec_{\overline{W}} M, z \prec_{\overline{W}} M)/2,$$

which is what we want. ∎

The final claim achieves our goal.

Claim 4. Property Ed(δ) is violated on W.

Proof. Suppose, for a contradiction, that the Ed(δ) constraint (5) is satisfied for W. Call $M \in \mathcal{M}$ *good* if there is at most one vertex of W which is not matched by M with another vertex of W, *i.e.*, if M uses $\lfloor |W|/2 \rfloor$ edges in W; otherwise, say M is *bad.* Writing $\mathcal{N}$ for the number of edges with both ends in W in a random (according to p) matching M, we have

$$(33) \qquad (1-\delta)\left\lfloor \frac{|W|}{2} \right\rfloor \geq \sum_{\substack{A \subseteq W \\ A \in \Gamma}} p_A = E[\mathcal{N}] \geq p(M \text{ good}) \left\lfloor \frac{|W|}{2} \right\rfloor$$

(the first inequality in (33) is (5), while the second follows from the definition of good). Thus (recall (25)), we have $p(M \text{ good}) \leq (1-\delta)$, or

$$(34) \qquad p(M \text{ bad}) \geq \delta.$$

If M is bad, then there are distinct $y, z \in W$, each of which is either unsaturated by M or matched by M with a vertex of $V \setminus W$. Thus,

$$p(M \text{ bad}) \leq$$

$$\sum_{\{y,z\} \in \binom{W}{2}} [p(y, z \not\sim M) + p(y \not\sim M, z \prec_{\overline{W}} M) + p(y \prec_{\overline{W}} M, z \not\sim M) + p(y, z \prec_{\overline{W}} M)].$$

Applying Claims 1–3 gives

$$(35) \qquad p(M \text{ bad}) \leq$$

$$\binom{|W|}{2}(\beta + 2\alpha^{-1} + 2\alpha^{-1} + 4\alpha^{-2}\beta^{-1}) \leq 2\delta^{-2}(\beta + 4\alpha^{-1} + 4\alpha^{-2}\beta^{-1}),$$

where we used $|W| \leq 2\delta^{-1}$ to eliminate the binomial coefficient. A routine computation, using the definitions of α, β (given following (23)) and the fact that $j_0 \geq 1$, shows that an appropriate choice of T (here $T \geq (\log_2 C\delta^{-5})^{1/2} + D$, where C, D are constants, suffices) forces the right side of (35) to be less than δ. Thus, $p(M \text{ bad}) < \delta$, which contradicts (34). ∎

Claim 4 provides the contradiction we set out to obtain in proving Lemma 8.1. ∎

Proof of Lemma 4.1

Form $\tilde{\Gamma}$ from Γ by replacing each $\{x\} \in \Gamma$ by a new 2-edge $\{x, x'\}$, where the x' are distinct new vertices. Then $\tilde{\Gamma}$ is 2-uniform on $\tilde{V} := V \mathring{\cup} \{x' : \{x\} \in \Gamma\}$. Define $\tilde{\lambda}: \tilde{\Gamma} \to \mathbf{R}^+$ by

$$\tilde{\lambda}(A) = \begin{cases} \lambda(A) & \text{if } A \in \Gamma_2 \\ \lambda(\{x\}) & \text{if } A = \{x, x'\}, \end{cases}$$

and let $\tilde{p}$ denote the resulting hard-core distribution on $\mathcal{M}(\tilde{\Gamma})$. Applying Lemma 8.1 to $\tilde{\Gamma}$, $\tilde{p}$ yields the desired δ'. ∎

Back to Lemma 4.3

We are almost ready to return to Lemma 4.3, which will follow as a simple corollary once we establish two results related to Lemma 4.1. The first is an analogue for triples while the second is a converse. We continue to let p denote a h.c.d. on $\mathcal{M}(\Gamma)$.

(8.2) Lemma. *For any $\delta \in (0,1)$ there is a $\delta'' = \delta''(\delta) \in (0,1)$ such that if p has property* Ed(δ), *then $p(\overline{x}, \overline{y}, \overline{z}) > \delta''$ for all $x, y, z \in V$.*

Proof. Using the device in the proof of Lemma 4.1, we may assume that Γ is a multigraph. The argument is then similar to those in the proofs of Claims 1–3 of Lemma 8.1. Generate a random M (according to p) by mating the random (according to p) and independently chosen M_1, M_2. Then

$$p(x, y, z \not\sim M) \geq \sum_{w \in V} p(x, y \not\sim M_1, \{z, w\} \in M_1) p(z, w \not\sim M_2) \frac{1}{8}. \tag{36}$$

Lemma 4.1 shows that each of the factors $p(z, w \not\sim M_2)$ on the right side of (36) is at least δ'. Thus, we have $p(x, y, z \not\sim M) \geq \delta' \sum_{w \in V} p(x, y \not\sim M_1, \{z, w\} \in M_1)/8$. But the last sum is simply $p(\overline{x}, \overline{y}) - p(\overline{x}, \overline{y}, \overline{z})$, so again applying Lemma 4.1, we find that $p(\overline{x}, \overline{y}, \overline{z})[1 + \delta'/8] > \delta'^2/8$, and the assertion follows. ∎

(8.3) Lemma. *For any $\delta \in (0,1)$ there is a $\delta' = \delta'(\delta) \in (0,1)$ such that if $p(\overline{x}, \overline{y}) > \delta$ for all $x, y \in V$, then p has property* Ed(δ').

(Note this is not the δ' of Lemma 4.1.)

Proof. We take $\delta' = \min\{\delta/2, \delta^2\}$ and verify (5) (with δ' in place of δ), (4) being immediate since $p(\overline{x}) \geq p(\overline{x}, \overline{y}) > \delta > \delta'$. For $|W|$ even, (5) follows from (4). For odd

$|W|>2/\delta$, we obtain (5) (in this case with $\delta/2$) from the fact that the $p(\overline{x})$'s are at least δ:

$$\sum_{\substack{F\subseteq W\\ F\in\Gamma_2}} p_F \le \frac{1}{2}\sum_{x\in W} p(x) = \frac{1}{2}\left(|W| - \sum_{x\in W} p(\overline{x})\right) \le \frac{(1-\delta)|W|}{2} \le (1-\delta/2)\left\lfloor\frac{|W|}{2}\right\rfloor.$$

It is only for odd $|W|\le 2/\delta$ that we need our full hypothesis. In this case we simply fix x,y (distinct) in W and note that

$$\sum_{\substack{F\subseteq W\\ F\in\Gamma_2}} p_F \le p(\overline{x},\overline{y})\left(\left\lfloor\frac{|W|}{2}\right\rfloor - 1\right) + (1-p(\overline{x},\overline{y}))\left\lfloor\frac{|W|}{2}\right\rfloor <$$

$$\left\lfloor\frac{|W|}{2}\right\rfloor - \delta < (1-\delta^2)\left\lfloor\frac{|W|}{2}\right\rfloor. \qquad \blacksquare$$

We at last give the proof of Lemma 4.3, whose statement is repeated for the readers' convenience.

(4.3) Lemma. *For any $\delta\in(0,1)$ there is a $\tilde\delta=\tilde\delta(\delta)\in(0,1)$ such that if p has property* Ed(δ), *then $p(\cdot|\overline{x})$ has property* Ed$(\tilde\delta)$ *for all $x\in V$.*

Proof. Lemma 8.2 provides a $\delta''=\delta''(\delta)>0$ such that for all $y,z\in V$, we have $p(\overline{y},\overline{z}|\overline{x})>\delta''$. Since $p(\cdot|\overline{x})$ is hard-core, Lemma 8.3 now shows that $\tilde\delta=\delta'\circ\delta''$ suffices (of course δ' here is the function guaranteed by Lemma 8.3). $\blacksquare$

Acknowledgement. The authors thank the referees for their careful reading of the manuscript and their thoughtful suggestions for improving the exposition.

References

[1] N. Alon and J. Spencer: *The Probabilistic Method*, Wiley, New York, 1992.

[2] J. van den Berg and J. E. Steif: Percolation and the hard-core lattice gas model, *Stochastic Process. Appl.*, **49** (1994), 179–197.

[3] C. Berge: *Hypergraphs: Combinatorics of Finite Sets*, North-Holland, Amsterdam, 1989.

[4] B. Bollobás: *Extremal Graph Theory*, Academic Press, London, 1978.

[5] J. Edmonds: Maximum matching and a polyhedron with 0,1-vertices, *J. Res. Nat. Bur. Standards (B)*, **69** (1965), 125–130.

[6] Z. Füredi: Matchings and covers in hypergraphs, *Graphs Combin.*, **4** (1988), 115–206.

[7] C. D. Godsil and I. Gutman: On the theory of the matching polynomial, *J. Graph Theory*, **5** (1981), 137–144.

[8] C. D. Godsil: *Algebraic Combinatorics*, Chapman and Hall, New York, 1993.

[9] O. J. Heilmann and E. H. Lieb: Monomers and dimers, *Phys. Rev. Letters*, **24** (1970), 1412–1414.

[10] O. J. Heilmann and E. H. Lieb: Theory of monomer-dimer systems, *Commun. Math. Phys.*, **25** (1972), 190–232.

[11] J. Kahn: Asymptotics of the chromatic index for multigraphs, *J. Combin. Theory Ser. B*, **68** (1996), 233–254.

[12] J. Kahn: Asymptotics of the list-chromatic index for multigraphs, submitted.

[13] J. Kahn: A normal law for matchings, submitted.

[14] J. Kahn and P. M. Kayll: Fractional v. integral covers in hypergraphs of bounded edge size, *J. Combin. Theory Ser. A*, **78** (1997), 199–235.

[15] J. Kahn and J.-H. Kim: Random matchings in regular graphs, *Combinatorica*, to appear.

[16] P. M. Kayll: *Asymptotically Good Covers in Hypergraphs*, Dissertation, Rutgers University, New Brunswick, NJ, 1994.

[17] P. M. Kayll: Asymptotically good covers in hypergraphs, *Diss. Summ. Math.*, **1** (1996), 9–16.

[18] P. M. Kayll: "Normal" distributions on matchings in a multigraph: overview with applications, *Congr. Numer.*, **107** (1995), 179–191.

[19] H. Kunz: Location of the zeros of the partition function for some classical lattice systems, *Phys. Lett.*, **32**A (1970), 311–312.

[20] C. W. Lee: Some recent results on convex polytopes, *Contemporary Mathematics*, **114** (1990), 3–19.

[21] C. W. Lee: Convex polytopes, the moment map, and canonical convex combinations, manuscript, 1994.

[22] G. M. Louth: *Stochastic Networks: Complexity, Dependence and Routing*, Dissertation, Cambridge University, 1990.

[23] L. Lovász and M. D. Plummer: *Matching Theory*, North-Holland, New York, 1986.

[24] J. R. Munkres: *Elements of Algebraic Topology*, Benjamin/Cummings, Menlo Park, 1984.

[25] Y. Rabinovich, A. Sinclair and A. Wigderson: Quadratic Dynamical Systems (Preliminary Version), in: *Proc. 33rd IEEE Symposium on Foundations of Computer Science*, pp. 304–313, 1992.

[26] A. Schrijver: *Theory of Linear and Integer Programming*, Wiley, New York, 1986.

[27] J. SPENCER: *Ten Lectures on the Probabilistic Method*, SIAM, Philadelphia, 1993.

Jeff Kahn

Department of Mathematics and RUTCOR
Rutgers University
New Brunswick, NJ 08903
U.S.A.
jkahn@math.rutgers.edu

P. Mark Kayll

Department of Mathematical Sciences
The University of Montana
Missoula, MT 59812-1032
U.S.A.
kayll@charlo.math.umt.edu