In view of previous results we now estimate the probability that $CM_n(d)$ is simple. We start with combinatorial result that we will not need later, but which is interesting itself.

(6.14) Theorem (Erdős–Gallai 1960). Let $d_1 \geq d_2 \geq \dots \geq d_n$ be such that $\sum_{i=1}^n d_i$ is even. There exists SIMPLE graph with the degree sequence $d_i$ iff

(6.15)
$$\sum_{i=1}^k d_i \leq k(k-1) + \sum_{i=k+1}^n \min(k, d_i).$$

Proof: Necessity of (6.15): The LHS of (6.15) is the total degree of the first $k$ vertices. The first term on the RHS is twice the maximal number of edges between those $k$ vertices, the second term on the RHS bounds the total number of edges between vertices in $[k]$ and vertices in $[n]\setminus[k]$.

Sufficiency of (6.15) Is harder, see references in [vdH, p 180]

This is the main result of this lecture.

(6.16) Theorem: Assume that $d^n$ satisfies (6.13 a–c). Then
$$\lim_{n\to\infty} P[CM_n(d^n) \text{ is simple}] = e^{-\frac{\nu}{2}-\frac{\nu^2}{4}}$$
where
$$\nu = \frac{E[D(D-1)]}{ED}.$$

The theorem is a direct corollary of the following proposition, giving stronger result. Let

$s_i = x_{ii}$ be the number of loops on $i \in [n]$

(6.17) $m_i = \sum_{j\neq i} (x_{ij}-1)_+$ be the # of additional multiple edges on $i$

$$S_n = \sum_{i\in[n]} s_i, \quad M_n = \sum_{i\in[n]} m_i$$

(6.18) <u>Proposition:</u> (6.13 a–c). The pair $(S_m, M_m)$ converges in distribution to $(S, M)$, where $S$ and $M$ are independent, $S = \mathrm{Pois}\left(\frac{\nu}{2}\right)$, $M = \mathrm{Pois}\left(\frac{\nu^2}{4}\right)$.

<u>Proof:</u> We will use

(6.19) <u>Theorem:</u> A vector $(X_1^n, \dots, X_k^n)_{n \geq 1}$ converges in distribution to a vector of independent Poissons $(X_1, \dots, X_k)$, $X_i \sim \mathrm{Pois}(\lambda_i)$ if for all $r_1, \dots, r_k \in \mathbb{N}$

$$\lim_{n \to \infty} E\left[ (X_1^n)_{r_1} \cdots (X_k^n)_{r_k} \right] = \lambda_1^{r_1} \cdots \lambda_k^{r_k}.$$

<u>Proof:</u> For $k = 1$ it was proved in exercises. $k > 1$ is a simple extension. $\square$

In view of (6.19). We must check

(6.20) $$\lim_{n \to \infty} E\left[ (S_n)_s (M_n)_r \right] = \left(\frac{\nu}{2}\right)^s \left(\frac{\nu^2}{4}\right)^r.$$

For $i \in [n]$, $1 \leq s < t \leq d_i$, let

(6.21) $$I_{st,i} = \mathbb{1}\{\text{half edge } (i,s) \text{ is paired with } (i,t)\}.$$

that is $\quad s_i = \sum_{1 \leq s < t \leq d_i} I_{st,i}$, and for $i,j \in [n]$, $s_1 < s_2 \leq d_i$, $t_1 \neq t_2 \leq d_j$

(6.22) $$I_{s_1t_1, s_2t_2, ij} = \mathbb{1}\{ (i, s_\ell) \text{ is paired to } (j, t_\ell), \ \ell = 1,2 \}.$$

We define

(6.23) $$\tilde{M}_n = \sum_{i<j} \sum_{s_1 < s_2 \leq d_i} \sum_{t_1 \neq t_2 \leq d_j} I_{s_1t_1, s_2t_2, ij}.$$

Observe that

(6.24) $$M_n \leq \tilde{M}_n \quad \text{and} \quad M_n = 0 \ \text{ iff } \ \tilde{M}_n = 0.$$

We will prove (6.18) with $M_n$ replaced by $\tilde{M}_n$. (6.18) then follows by

(6.25) <u>Exercise:</u> Show that $\lim P(M_n \neq \tilde{M}_n) = 0$.

We set $\mathcal{I}_1 = \{ (st, i) : i \in [n], s < t \leq d_i \}$, $\mathcal{I}_2 = \{ (s_1t_1, s_2t_2, ij) : i,j \in [n], s_1 < s_2 \leq d_i, t_1 \neq t_2 \leq d_j \}$

and write for $m \in \mathcal{I}_1$, $I_m = I_{st,i}$ if $m = (st, i)$, and for $m = (s_1t_1, s_2t_2, ij) \in \mathcal{I}_2$ we set $I_m = I_{s_1t_1, s_2t_2, ij}$.

We know from exercises that

(6.26) $$E[(S_n)_s (\tilde{R}_n)_r] = \sum^*_{\substack{m_1,\dots,m_s \in \mathcal{I}_1 \\ \bar{m}_1,\dots,\bar{m}_r \in \mathcal{I}_2}} P(I_{m_1} = \dots = I_{m_s} = I_{\bar{m}_1} = \dots = I_{\bar{m}_r} = 1),$$

where $\Sigma^*$ is a sum over distinct indices. Since half edges are paired uniformly

(6.27) $$P(I_{m_1} = \dots = I_{\bar{m}_r} = 1) = \prod_{i=0}^{s+2r-1} (\ell_n - 1 - 2i)^{-1},$$

unless there is a conflict between pairings required by $m$'s and $\bar{m}$'s, in which case $P(I_{m_1} = \dots = I_{\bar{m}_r} = 1) = 0$.

Hence,

(6.28) $$E[(S_n)_s (\tilde{R}_n)_r] \le \sum^*_{\substack{m_1,\dots,m_s \in \mathcal{I}_1 \\ \bar{m}_1,\dots,\bar{m}_r \in \mathcal{I}_2}} \prod_{i=0}^{s+2r-1} (\ell_n - 1 - 2i)^{-1} = \frac{|\mathcal{I}_1| \cdots (|\mathcal{I}_1| - s + 1)|\mathcal{I}_2| \cdots (|\mathcal{I}_2| - r + 1)}{(\ell_n - 1) \cdots (\ell_n - 2s - 4r + 1)}$$

W.l.o.g. we may assume that $|\mathcal{I}_1|, |\mathcal{I}_2| \nearrow \infty$ since otherwise $n = n_1(1 - o(1))$ and then $(S_n, \tilde{R}_n) = 0$ whp and also $\nu = 0$.

(6.29). Hence, $$(6.28) \le \left(\lim_{n\to\infty} \frac{|\mathcal{I}_1|}{\ell_n}\right)^s \left(\lim_{n\to\infty} \frac{|\mathcal{I}_2|}{\ell_n^2}\right)^r.$$

By (6.13 a-c)

$$\lim_{n\to\infty} \frac{|\mathcal{I}_1|}{\ell_n} = \lim_{n\to\infty} \frac{1}{\ell_n} \sum_{i\in[n]} \frac{d_i(d_i-1)}{2} = \lim \frac{1}{ED_n} E\left[\frac{D_n(D_n-1)}{2}\right] = \frac{\nu}{2}$$

and

$$\lim_{n\to\infty} \frac{|\mathcal{I}_2|}{\ell_n^2} = \lim_{n\to\infty} \frac{1}{\ell_n^2} \sum_{i<j\in[n]} \frac{d_i(d_i-1)}{2} d_j(d_j-1) =$$

$$= \left(\lim_{n\to\infty} \frac{1}{\ell_n} \sum_{i\in[n]} \frac{d_i(d_i-1)}{2}\right)^2 - \lim_{n\to\infty} \frac{1}{\ell_n^2} \sum \frac{d_i^2(d_i-1)^2}{2} =$$

$$= \left(\frac{\nu}{2}\right)^2.$$

(6.30) Exercise: Prove that (6.13) implies $\max_{i\in[n]} d_i = o(\sqrt{n})$ and use it to prove the last statement in more detail.

This proves the upper bound for (6.20) and completes the proof for $\nu = 0$.

For a lower bound in the case $\nu > 0$, we observe that

$$\sum^*_{\substack{m_1,\dots,m_s\in\mathcal{I}_1\\ \bar m_1\dots\bar m_r\in\mathcal{I}_2}} \prod_{i=0}^{s+2r-1}(\ell_n-1-2i)^{-1} - E\left[(S_n)_s(\tilde{R}_n)_r\right] =$$

(6.26) $$= \sum^*_{\substack{m_1,\dots,m_s\\ \bar m_1,\dots,\bar m_r}} \left(J_{m_1,\dots,m_s,\bar m_1,\dots,\bar m_r}\right) \prod_{i=0}^{s+2r-1}(\ell_n-1-2i)^{-1},$$

where $J_{m_1,\dots,m_s,\bar m_1,\dots,\bar m_r}$ is the indicator of conflict between indices $m_1,\dots,\bar m_r$. We should show that the RHS of (6.26) vanishes as $n\nearrow\infty$. There are 3 cases:

– conflict between $m_i$'s: Note that $m_{i_a}=(st,i)$ conflicts with $m_{i_b}=(s't',j)$ iff $i=j$ and at least two of $s,t,s',t'$ agree the number of such conflicting pairs is at most

(6.27) $$\sum_{i\in[n]} 6d_i^3 = o\Big(\sum_{i\in[n]} d_i(d_i-1)\Big)^2 = o(|\mathcal{I}_1|^2),$$

where we used (6.30), $\nu=\lim \frac{2|\mathcal{I}_1|}{\ell_n}>0$. The contribution of such conflicts to (6.26) is thus smaller than

(6.28) $$|\mathcal{I}_1|^{s-2}|\mathcal{I}_2|^r\prod_{i=1}^{s+2r-1}(\ell_n-1-2i)^{-1}\cdot o(|\mathcal{I}_1|^2)\longrightarrow 0,$$

comparing this with (6.28),…,(6.29).

– conflict between $m_{i_a}$ and $\bar m_{i_b}$. $m_{i_a}=(st,i)$ conflicts with $\bar m_{i_b}=(s_1t_1,s_2t_2,kl)$ if $i=k$ and at least two of $s,t,s_1,t_1$ agree or $i=l$ and at least two of $s,t,s_2,t_2$ agree. The number of the conflicting pairs is at most

(6.29) $$6\cdot\sum_{i\in[n]} d_i^3\sum_{j\in[n]} d_j^2 = o\Big(\sum_{i\in[n]} d_i(d_i-1)\Big)^3 = o(|\mathcal{I}_1||\mathcal{I}_2|),$$

and by the same argument as in the first case, we see that the contribution in this case vanishes.

– conflict between $\bar m_{i_a}=(s_1t_1,s_2t_2,ij)$ and $\bar m_{i_b}=(s_1't_1',s_2't_2',i'j')$ occurs if at least two of $i,j,i',j'$ agree, say $i,i'$ and then at least two of $s,t,s',t'$ agree. We obtain a bound

(6.30) $$\le 6\sum_{i\in[n]} d_i^3\sum_{j\in[n]}d_j^2\sum_{k\in[n]}d_k^2 = o\Big(\sum_{i\in[n]} d_i(d_i-1)\Big)^4 = o(|\mathcal{I}_2|^2)$$

□

Combining (6.16) and (6.7), (6.9) we obtain

(6.31) <u>Corollary</u>: Assume (6.13 (a-c)). The number of simple graphs with degree sequence $(d_i)_{i\in[n]}$ is

$$e^{-\nu/2-\nu^2/4}\,\frac{(\ell_n-1)!!}{\prod_{i\in[n]} d_i!}\,(1+o(1))$$

In particular, if $d_i = r$ $\forall i\in[n]$ ($r$-regular graph), this is

$$e^{-\frac{(r-1)}{2}-\frac{(r-1)^2}{4}}\,\frac{(rn-1)!!}{(r!)^n}\,(1+o(1))$$

We have also the following useful equivalence.

(6.32) <u>Corollary</u>: (6.13 a-c) If event(s) $\mathcal{E}_n$ occur with high probability for $CM_n(d)$, then they occur w.h.p for uniform simple graph with degree sequence $d$.

## Component structure in Configuration model

We now explore the connectivity structure in the configuration model. The key result is Theorem (6.34) below, providing a local approximation by certain random tree.

The random tree is defined as follows. Assume

(6.33) that (6.13) hold and let $p_k = P(D=k)$, $k \geq 0$.

Define now another distribution

$$p_k^* = \frac{(k+1)\, p_{k+1}}{ED}, \quad k \geq 0$$

which is easily proved to statisfy $\sum_{k\geq 0} p_k^* = 1$.

We consider rooted random tree, denoted BP, where the root have $p_k$-distributed number of offsprings and all other vertices have $p_k^*$-distributed number of offsprings.

We also write $BP(t)$, $t \in \mathbb{N}$, for the subtree of BP obtained by the first $t$ steps of breadth-first exploration. Finally let $G_n(t)$, $t \in \mathbb{N}$, denotes the graph obtained by $t$-steps of breadth-first exploration on $CM_n(d)$ started from uniformly chosen vertex $U_n$.

(6.34) <u>Theorem</u>: Ass. (6.13), there is a coupling $(\hat{G}_n(t), \hat{BP}(t))_{t\geq 0}$ of $(G_n(t))_{t\geq 0}$ and $(BP(t))_{t\geq 0}$ such that

$$\mathbb{P}\left[ (\hat{G}_n(t))_{t\leq m} \neq (\hat{BP}(t))_{t\leq m} \right] \xrightarrow{n\to\infty} 0,$$

for every fixed $m$.

(6.35) <u>Remark</u>: One can even assume $m \nearrow \infty$ sufficiently slowly.

<u>Proof:</u> For the root, the degree of $U$ in $CM_n(d)$ is $D_n$. As $D_n \to D$, it is possible to couple $D_n$ with $D$ s.t. $P(D_n \neq D) \to 0$. So the coupling works for "$t=0$".

Assume now that we constructed $(\hat{G}_n(t), \hat{BP}(t))_{t\leq k-1}$.

To obtain $\hat{G}_n(t)$ for $t=k$, we take first unpaired half edge $x_k$ and pair it to uniform half edge $y_k$ not paired so far. Connect $y_k$ with $x_k$ to create $\hat{G}_n(k)$. Observe that if $v_k$ is the vertex attached to $y_k$, then the new attached vertex has $d_{v_k}-1$ offsprings, if we connected really new vertex (which has a large probability as $n\nearrow\infty$). Moreover

$$P(d_{v_k}=l,\ v_k \text{ is new}) = \frac{\sum_{v\notin G_{k-1},\, d_{v_k}=l} l}{\#\text{ unused half edges after } k-1 \text{ steps}}$$

$$\simeq \frac{n\cdot p_l\cdot l}{\sum_{l'\geq 0} n p_{l'} l'} = p^*_{l+1}.$$ □.

One can obtain even more precise results in this direction, see e.g. vdH2, Ch 3.

(6.36) <u>Lemma</u> (phase transition for BP).

$P(|BP|=\infty)>0$ iff $E[D^*] = \frac{E D(D-1)}{ED} = \nu > 1$ exercise!

$D^*$ is $p^*$-distributed.

<u>Proof:</u> This is an easy extension of (2.4).

As consequence of (6.36) and (6.39) it is plausible to believe that the following is true. For proofs see [vdH2].

(6.37) <u>Theorem</u> (phase transition in CM). Ass. (6.13) and $p_2<1$.

(a) If $\nu>1$, then $\frac{|C_{max}|}{n} \xrightarrow{P} \xi \in (0,1]$, and $\frac{|C^{(2)}|}{n} \xrightarrow{P} 0$.

(b) If $\nu<1$, then $|C_{max}| \leq C \lg n$ whp.

(6.38) <u>Remark:</u> Even for $\nu=1$, the behaviour is as in ER case.