Fast sequential algorithm for generating directed random graphs with a given degree sequence

Femke van Ieperen[1] and Ivan Kryven[1,*]

[1]Mathematical Institute, Utrecht University, PO Box 80010, 3508 TA Utrecht, the Netherlands
[*]i.kryven@uu.nl

Abstract

We propose a near-linear complexity algorithm for generating simple directed random graphs with a given degree sequence and show that this algorithm provides a means of uniform sampling for large graphs. The algorithm is applicable when the maximum degree, $d_{\max}$, is asymptotically dominated by $m^{1/4}$ with m being the number of edges and admits an implementation with the expected running time of the order of $md_{\max}$.
Keywords: Random Graphs, Directed Graphs, Randomised Approximation Algorithms
MSC: 05C80, 05C20, 68W20, 68W25

A graph is simple when it has no multi edges or self loops. Given a graphical degree sequence, the existence of a simple graph that features this sequence is guaranteed by the Erdős-Gallai theorem. However, enforcing this property when uniformly sampling from the set of all graphs that satisfy a given degree sequence is not a straightforward task. One computationally expensive way is the rejection sampling with configurational model [1]. According to this method a multigraph is constructed by randomly matching half edges of a given graphical degree sequence. Repeating this construction multiple times will eventually produce a simple graph in time that is linear in the number of vertices but exponential in the square of the average degree. This strategy can be improved if instead of rejecting every multigraph, we fix those multigraphs that are not too bad by switching several edges to remove edge multiplicity and loops. Such a switching procedure was shown to implement exact uniform sampling in polynomial time for regular graphs by McKay and Wormald [2] and more general degree sequences by Gao and Wormald [3, 4].

Generating random graphs is closely related to counting and generation of binary matrices with given properties, such as row and column sums, which are of general interest to combinatorialists, statisticians, and computer scientists. Uniform generation of simple graphs is used in analysis of algorithms and networks [5, 6, 7]. In algorithmic spectral graph theory, fast sampling is required to study spectra of sparse random matrices [8, 9, 10], where beyond the case of undirected graphs, heuristic algorithms had to be postulated. Much research has therefore been directed towards sampling more general simple graphs with fast but only asymptotically exact methods, which concentrate mainly on two ideas: 1) Markov Chain Monte Carlo (MCMC) algorithms approximate the desired sample by taking the last element of an ergodic Markov chain [11, 12, 13, 22, 24]; the uniformity is asymptotically exact for long chains and any finite number of nodes but the performance depends on the initial seed and have super-linear complexity. 2) Fast sequential construction algorithms [14, 15, 16] build a graph by placing m edges one-by-one, starting with an empty graph; they typically feature almost linear complexity but the uniformity is asymptotically exact only for large graphs.

For a given number of nodes n, one can always improve the expected error bound on the output distribution in an MCMC method by running the algorithm longer. However, showing that this sample is sufficiently independent from the initial seed, *i.e.* estimating the mixing time of the chain, has been achieved only of several classes of random graphs. First MCMC algorithms were proposed for random regular graphs [11, 17, 18, 12], and were also shown to feature polynomial mixing time [19, 20]. Later, switching MCMC algorithms were suggested for graphs with arbitrary degree sequences by Kannan, Tetali and Vempala [21], wherein the rapid mixing property was shown by P.L. Erdős et al. [22] for the class of P-stable [23] degree distributions, and more recently, for other stability classes by Gao and Greenhill [24]. See also, Jason [25] for the analysis of the convergence to uniformity. Bergerand and Müller-Hannemann suggested a MCMC algorithm for sampling random digraphs [13], with some relevant rapid mixing results shown by Greenhill [26, 27] and P.L. Erdős et al. [28]. Further generalisations were also proposed for degree-correlated random graphs [29, 30, 29]. There are also approaches [31, 32] that realise non-uniform sampling while also outputing probability of the generated sample a posteriori. Hence they may be used to compute expectations over the probability space of random graphs.

As an alternative to MCMC, linear complexity sequential algorithms construct simple graphs by starting with an empty edge set and adding edges one-by-one while updating the probability after each edge placement [14, 33, 34, 15]. For instance, Steger and Wormald's algorithm [14] samples regular graphs almost uniformly with the running time shown to be $\mathcal{O}(nd_{\max}^2)$ by Kim and Vu [34]. Bayati, Kim and Saberi [15] generalised the sequential method to an arbitrary degree sequence, yet maintaining a near-linear in the number of edges algorithmic complexity. For these algorithms, the maximum degree may depend on n with some asymptotic constraints, and the bounds on the error in the output distribution asymptotically vanish as n tends to infinity. The downside of sequential algorithms is that for a fixed n they alone cannot improve the error by performing more computations. Nevertheless, the error can be still improved by generating a seed with a sequential algorithm and then post-processing with an MCMC to improve uniformity, providing that the chain features rapid mixing for the given degree distribution.

In this work, we provide a fast sequential algorithm for almost-unifrom sampling of simple directed graphs with a given degree sequence by building upon the work of Bayati, Kim and Saberi [15]. We call a digraph *simple* when it has no self loops or parallel edges with identical orientation. Similarly to the undirected case, our degree sequence is required to be graphical in the sense of [35]. The expected runtime of our algorithm is almost-linear in the number of edges and the bound on the error between the uniform and output distributions asymptotically vanishes for large graphs. As such, our algorithm provides a good trade-off between the speed and uniformity. Furthermore, if exact uniformity is required for finite n, it can be achieved by post-processing with an appropriate MCMC algorithm for directed graphs, see for example [27, 28].

We explain the algorithm in Section 1. The proof that this algorithm generates graphs distributed within up to a factor of $1 \pm o(1)$ of uniformity is presented in Section 2 and is inspired by the proof of Bayati, Kim and Saberi [15, Section 7], wherein Vu's concentration inequality [36] plays a significant role. Our algorithm may fail to construct a graph, but it is shown that this happens with probability $o(1)$ in Section 3, following on [15, Section 5]. This work is completed with the runtime analysis of the algorithm in Section 4.

1 The algorithm

The algorithm is best explained as a modification of the directed configuration model, which generates a configuration by sequentially matching a random in-stub to a random out-stub. One can therefore see that generating a uniformly random configuration is not difficult, however, a random configuration may induce a multigraph, which we do not desire. This issue can be remedied by the following procedure: a match between the chosen in- and out-stub is rejected if it leads to a self-loop or multi-edge. Then, the resulting configuration necessarily induces a simple graph. Note that this rejection of specific matches destroys the uniformity of the generated graphs. To cancel out the non-uniformity bias, we accept each admissible match between an in- and an out-stub with a cleverly chosen probability, which restores the uniformity of the samples. Namely, we show that the distribution of the resulting graphs is within $1 \pm o(1)$ of uniformity for large graphs. Another consequence of the constraint on acceptable matches, is that it may result in a failed attempt to finish a configuration, for example, if at some step of the matching procedure the only remaining stubs consist of one in-stub and one out-stub belonging to the same vertex. In this case, we reject the entire configuration and start from scratch again. As we will show later in Section 4, a failure is not likely to occur, *i.e.* the probability that a configuration cannot be finished is $o(1)$.

Algorithm 1: generating simple directed graphs obeying a given degree sequence

Input : $\mathbf{d}$, a graphical degree sequence without isolated nodes
Output: $G_{\mathbf{d}} = (V, E)$ a digraph obeying $\mathbf{d}$ and N an estimation for the number of simple digraphs obeying $\mathbf{d}$ or a failure

$V = \{1, 2, \dots n\}$ `// set of vertices`
$\hat{d} = \mathbf{d}$ `// residual degree`
$E = \emptyset$ `// set of edges`
$P = 1$ `// probability of generating this ordering`
while *edges can be added to E* **do**
  Pick $i, j \in V$ with probability P_{ij} proportional to $\hat{d}_i^+ \hat{d}_j^- \left(1 - \frac{d_i^+ d_j^-}{2m}\right)$ amongst all ordered pairs (i, j) with $i \neq j$ and $(i, j) \notin E$;
  Add (i, j) to E, decrease $\hat{d}_i^+$ and $\hat{d}_j^-$ by 1 and set $P = P \cdot P_{ij}$;
if $|E| = m$ **then**
  Return $G_{\mathbf{d}} = (V, E)$, $N = \frac{1}{m!P}$
else
  Return `failure`

The pseudo-code of our algorithm, shown in Algorithm 1, is a generalisation of [15, Procedure A], written for undirected graphs. We use the following notation: Let $\mathbf{d} = \{(d_i^-, d_i^+)\}_{i=1}^n$ with $d_i^-, d_i^+ \in \mathbb{N}$ be a graphical degree sequence, and $m = \sum_{i>0} d_i^- = \sum_{i>0} d_i^+$ the total number of edges. Furthermore we define

$$d_{\max} = \max\{\max\{d_1^-, d_2^-, \ldots, d_n^-\}, \max\{d_1^+, d_2^+, \ldots, d_n^+\}\}.$$

We wish to construct a simple directed graph $G_{\mathbf{d}} = (V, E)$ with vertex set $V = \{1, \ldots, n\}$ and edge set E that satisfies $\mathbf{d}$. At each step, Algorithm 1 chooses edge (i, j) with probability

$$P_{ij} \sim \begin{cases} \hat{d}_i^+ \hat{d}_j^- \left(1 - \frac{d_i^+ d_j^-}{2m}\right), & i \neq j \text{ and } (i,j) \notin E, \\ 0, & i = j \text{ or } (i,j) \in E, \end{cases}$$

and adds it to E, where the residual in-degree $\hat{d}_i^-$ (respectively residual out-degree $\hat{d}_i^+$) of vertex i is the number of unmatched in-stubs (out-stubs) of this vertex and E the set of edges constructed so far. When for all pairs $i, j \in V$ with $\hat{d}_i^+ > 0$ and $\hat{d}_j^- > 0$ there holds $i = j$ or $(i, j) \in E$, no edge can be added to E and the algorithm terminates. If the algorithm terminates before m edges have been added to E, it has failed to construct a simple graph obeying the desired degree sequence and outputs a *failure*. If the algorithm terminates with $|E| = m$, it returns a simple graph that obeys the degree sequence $\mathbf{d}$. In this case the algorithm also computes the total probability P of constructing the instance of $G_{\mathbf{d}}$ in the order it has been constructed. We will show that asymptotically each ordering of a set of m edges is generated with the same probability. Hence, the probability that the algorithm generates digraph $G_{\mathbf{d}}$ is asymptotically $m!P$. We will also show that each digraph is generated within a factor of $1 \pm o(1)$ of uniformity, and therefore $N = \frac{1}{m!P}$ is an approximation to the number of simple digraphs obeying the degree sequence. The value of N is also returned by the algorithm if it successfully terminates. To make these statements more precise, let us consider *degree progression* $\{\mathbf{d}_n\}_{n\in\mathbb{N}}$, that is a sequence of degree sequences indexed by the number of vertices n. The algorithm has the following favourable properties.

Theorem 1.1. *Let all degree sequences in $\{\mathbf{d}_n\}_{n\in\mathbb{N}}$ are graphical and such that for some $\tau > 0$, the maximum degree $d_{\max,n} = \mathcal{O}\left(m^{1/4-\tau}\right)$, where m is the number of edges in $\mathbf{d}_n$. Then Algorithm 1 applied to $\mathbf{d}_n$ terminates successfully with probability $1 - o(1)$ and has an expected runtime of $\mathcal{O}(m d_{\max})$. Furthermore, the output graph $G_{\mathbf{d}_n}$ is generated with a probability within factor $1 \pm o(1)$ of uniformity.*

The remainder of this work covers the proof of Theorem 1.1, which is split into three parts discussing the uniformity of the generated digraphs, the failure probability of the algorithm and its runtime.

2 The probability that Algorithm 1 generates a given digraph

If the algorithm successfully terminates, the output graph $G_{\mathbf{d}}$ satisfies the desired graphical degree sequence $\mathbf{d}$ by construction. This section is devoted to showing that any $G_{\mathbf{d}}$ is generated with a probability within $1 \pm o(1)$ of the uniform probability. More formally we prove the following theorem:

Theorem 2.1. *Let $\mathbf{d}$ be a graphical degree sequence with maximum degree $d_{\max} = \mathcal{O}\left(m^{1/4-\tau}\right)$ for some $\tau > 0$. Let $G_{\mathbf{d}}$ be a random simple graph obeying this degree sequence. Then*

Algorithm 1 generates $G_{\mathbf{d}}$ with probability

$$[1+o(1)]\left(\frac{m!}{\prod_{r=0}^{m-1}(m-r)^2}\prod_{i=1}^{n}d_i^+!\prod_{i=1}^{n}d_i^-!\,e^{\frac{\sum_{i=1}^{n}d_i^-d_i^+}{m}-\frac{\sum_{i=1}^{n}(d_i^-)^2+(d_i^+)^2}{2m}+\frac{\sum_{i=1}^{n}(d_i^-)^2\sum_{i=1}^{n}(d_i^+)^2}{4m^2}+\frac{1}{2}}\right).$$

The proof is split into four steps, Sections 2.1-2.4. In Section 2.1, we start with determining the probability that the algorithm generates a given digraph $G_{\mathbf{d}}$.

2.1 The probability of generating a given digraph $G_{\mathbf{d}}$

Our goal is to determine the probability $\mathbb{P}_A(G_{\mathbf{d}})$ that Algorithm 1 outputs a given digraph $G_{\mathbf{d}}$ on input of a graphical $\mathbf{d}$. The output of Algorithm 1 can be viewed as a configuration in the following sense.

Definition 2.2. Let $\mathbf{d}$ be a degree sequence. For all $i \in \{1,2,\dots,n\}$ define a set of *in-stubs* W_i^- consisting of d_i^- unique elements and a set *out-stub* W_i^+ containing d_i^+ elements. Let $W^- = \cup_{i\in\{1,2,\dots,n\}}W_i^-$ and $W^+ = \cup_{i\in\{1,2,\dots,n\}}W_i^+$. Then a *configuration* is a random perfect bipartite matching of W^- and W^+, that is a set of tuples (a,b) such that each tuple contains one element from W^- and one from W^+ and each element of W^- and W^+ appears in exactly one tuple.

A configuration $\mathcal{M}$ prescribes a matching for all stubs, and therefore, defines a multigraph with vertices $V = \{1,2,\dots,n\}$ and edge set

$$E = [(i,j) \mid W_i^+ \ni a, W_j^- \ni b, \text{ and } (a,b) \in \mathcal{M}]. \tag{1}$$

Remark that the output of Algorithm 1 can be viewed as a configuration since at each step an edge (i,j) is chosen with probability proportional to $\hat{d}_i^+\hat{d}_j^-$, *i.e.* the number of pairs of unmatched out-stubs of i with unmatched in-stubs of j. Let $R(G_{\mathbf{d}}) = \{\mathcal{M} \mid G_{\mathcal{M}} = G_{\mathbf{d}}\}$ be the set of all configurations on (W^-,W^+) that correspond to $G_{\mathbf{d}}$. Since the output of Algorithm 1 is a configuration, there holds

$$\mathbb{P}_A(G_{\mathbf{d}}) = \sum_{\mathcal{M}\in R(G_{\mathbf{d}})}\mathbb{P}_A(\mathcal{M}).$$

Different configurations correspond to the same graph if they differ only in the labelling of the stubs. Since the algorithm chooses stubs without any particular order preference, each configuration in $R(G_{\mathbf{d}})$ is generated with equal probability. However, the probability to match an out-stub of i to an in-stub of j at a given step of the algorithm depends on the partial configuration constructed so far. Hence the order in which the matches are chosen, influences the probability of generating a configuration $\mathcal{M}$. Let for a given $\mathcal{M} \in R(G_{\mathbf{d}})$, $S(\mathcal{M})$ be the set of all the orderings $\mathcal{N}$ in which the configuration can be created. Because the configuration already determines the match for each in-stub, an ordering of $\mathcal{M}$ can be thought of as an enumeration of edges $\mathcal{N} = (e_1,e_2,\dots,e_m)$, $e_i \in E$, defining which in-stub gets matched first, which second, etc. There are $m!$ different orderings of the configuration $\mathcal{M}$. This implies that

$$\mathbb{P}_A(G_{\mathbf{d}}) = \prod_{i=1}^{n}d_i^-!\prod_{i=1}^{n}d_i^+!\sum_{\mathcal{N}\in S(\mathcal{M})}\mathbb{P}_A(\mathcal{N}).$$

Hence, we further investigate $\mathbb{P}_A(\mathcal{N})$. If the algorithm has constructed the first r elements of $\mathcal{N}$, it is said to be at step $r \in \{0, 1, \ldots, m-1\}$. There is no step m, as the algorithm terminates immediately after constructing the m^{th} edge. Let $d_i^{-(r)}$ (respectively $d_i^{+(r)}$) denote the number of unmatched in-stubs (out-stubs) of the vertex i at step r. Let E_r be the set of admissible edges that can be added to the ordering at step r,

$$E_r := \left\{(i,j) \mid i,j \in V,\ d_i^{+(r)} > 0,\ d_j^{-(r)} > 0,\ i \neq j,\ (i,j) \notin \{e_1, e_2 \ldots, e_r\}\right\}.$$

With this notation in mind, we write the probability of generating the entire ordering $\mathcal{N}$ as

$$\mathbb{P}_A(\mathcal{N}) = \prod_{r=0}^{m-1} \mathbb{P}\left[e_{r+1} | e_1, \ldots, e_r\right],$$

where

$$\mathbb{P}\left[e_{r+1} = (i,j) | e_1, \ldots, e_r\right] = \frac{1 - \frac{d_i^+ d_j^-}{2m}}{\sum_{(u,v)\in E_r} d_u^{+(r)} d_v^{-(r)} \left(1 - \frac{d_u^+ d_v^-}{2m}\right)}.$$

Here we slightly abuse the notation as this is the conditional probability that a given out-stub of i is matched with a given in-stub of j, rather than the conditional probability that the edge (i,j) is created. The probability that the algorithm generates the graph $G_{\mathbf{d}}$ can be written as

$$\mathbb{P}_A(G_{\mathbf{d}}) = \prod_{i=1}^{n} d_i^-! \prod_{i=1}^{n} d_i^+! \prod_{(i,j)\in G_{\mathbf{d}}} \left(1 - \frac{d_i^+ d_j^-}{2m}\right) \sum_{\mathcal{N}\in S(\mathcal{M})} \prod_{r=0}^{m-1} \frac{1}{(m-r)^2 - \Psi_r(\mathcal{N})}, \tag{2}$$

where

$$\Psi_r(\mathcal{N}) = \sum_{(u,v)\notin E_r} d_u^{+(r)} d_v^{-(r)} + \sum_{(u,v)\in E_r} d_u^{+(r)} d_v^{-(r)} \frac{d_u^+ d_v^-}{2m}. \tag{3}$$

By comparing the expression (2) with the statement of Theorem 2.1, we observe that the proof will be completed if we show that for some ψ_r, which we define in section 2.2, there holds

$$\sum_{\mathcal{N}\in S(\mathcal{M})} \prod_{r=0}^{m-1} \frac{1}{(m-r)^2 - \Psi_r(\mathcal{N})} = [1 + o(1)]\, m! \prod_{r=0}^{m-1} \frac{1}{(m-r)^2 - \psi_r}, \tag{4}$$

and

$$\prod_{r=0}^{m-1} \frac{1}{(m-r)^2 - \psi_r} = [1 + o(1)] \prod_{r=0}^{m-1} \frac{1}{(m-r)^2} e^{\frac{\sum_{i=1}^n d_i^- d_i^+}{m} - \frac{\sum_{i=1}^n (d_i^-)^2 + (d_i^+)^2}{2m} + \frac{\sum_{i=1}^n (d_i^-)^2 \sum_{i=1}^n (d_i^+)^2}{4m^2}} \times e^{\frac{\sum_{(i,j)\in G_{\mathbf{d}}} d_i^+ d_j^-}{2m} + \frac{1}{2}}. \tag{5}$$

Indeed, combining the latter two equations with (2) and using that $1 - x = e^{-x + \mathcal{O}(x^2)}$ we find:

$$\mathbb{P}_A(G_{\mathbf{d}}) = [1 + o(1)]\, m! \prod_{i=1}^n d_i^-! \prod_{i=1}^n d_i^+! \prod_{r=0}^{m-1} \frac{1}{(m-r)^2} e^{\frac{\sum_{i=1}^n d_i^- d_i^+}{m} - \frac{\sum_{i=1}^n (d_i^-)^2 + (d_i^+)^2}{2m} + \frac{\sum_{i=1}^n (d_i^-)^2 \sum_{i=1}^n (d_i^+)^2}{4m^2} + \frac{1}{2}},$$

which coincides with the statement of Theorem 2.1. Thus proving equations (4) and (5) suffices to show validity of Theorem 2.1.

2.2 Defining ψ_r

The quantity $\Psi_r(\mathcal{N})$ is defined as a function of $\mathcal{N}$. We abbreviate $\Psi_r(\mathcal{N})$ by Ψ_r whenever $\mathcal{N}$ follows from the context. It can also be viewed as a function on the subgraph of $G_{\mathbf{d}}$ induced by the first r elements the ordering $\mathcal{N}$, which we denote by $G_{\mathcal{N}_r}$. Hence, when taking the expected value of Ψ_r over all orderings, we look at a random subgraph of $G_{\mathbf{d}}$ with exactly r edges. Closely related to this is the G_{p_r} model where we take a subgraph of $G_{\mathbf{d}}$ where each edge is present with probability $p_r = \frac{r}{m}$. We define $\psi_r = \mathbb{E}_{p_r}[\Psi_r]$. In the remainder of this section we will determine the value of ψ_r. To achieve this, let us first have a closer look at Ψ_r.

We split Ψ_r into a sum of two terms:

$$\Psi_r = \Delta_r + \Lambda_r,$$

with

$$\Delta_r = \sum_{(u,v)\notin E_r} d_u^{+(r)} d_v^{-(r)} \quad \text{and} \quad \Lambda_r = \sum_{(u,v)\in E_r} d_u^{+(r)} d_v^{-(r)} \frac{d_u^+ d_v^-}{2m}. \tag{6}$$

Note that Δ_r counts the number of *unsuitable pairs*, i.e. the number of pairs of the unmatched in-stubs with out-stubs that induce a self-loop or multi-edge. In the sequel we refer to a combination of an unmatched in- and out-stub as a *pair*. To simplify the analysis of Δ_r and Λ_r, they are also written as a sum of several quantities. Hence we further split

$$\Delta_r = \Delta_r^1 + \Delta_r^2$$

where

$$\Delta_r^1 = \sum_{i=1}^{n} d_i^{-(r)} d_i^{+(r)} \tag{7}$$

is the number pairs creating a self-loop, and

$$\Delta_r^2 = \Delta_r - \Delta_r^1, \tag{8}$$

is the number of pairs creating a double edge. The quantity Λ_r is also split into two terms:

$$\Lambda_r = \frac{\Lambda_r^{1+}\Lambda_r^{1-} - \Lambda_r^2}{4m} - \frac{\Lambda_r^3}{2m}, \tag{9}$$

with

$$\Lambda_r^{1+} = \sum_{i=1}^{n} d_i^{+(r)} d_i^+, \quad \Lambda_r^{1-} = \sum_{i=1}^{n} d_i^{-(r)} d_i^-, \tag{10}$$

$$\Lambda_r^2 = \sum_{i=1}^{n} d_i^{+(r)} d_i^+ d_i^{-(r)} d_i^- \quad \text{and} \tag{11}$$

$$\Lambda_r^3 = \sum_{\substack{(u,v)\notin E_r \\ u\neq v}} d_u^{+(r)} d_v^{-(r)} d_u^+ d_v^-. \tag{12}$$

We will now derive several bounds on the latter quantities, to be used in Section 2.4.

Lemma 2.3. *For all* $0 \le r \le m-1$ *there holds:*

(i) $\Delta_r \le (m-r)d_{\max}^2$*;*

(ii) $\Lambda_r^{1+} \le d_{\max}(m-r)$, $\Lambda_r^{1-} \le d_{\max}(m-r)$*;*

(iii) $\Lambda_r \le \frac{d_{\max}^2}{2m}(m-r)^2$.

Proof. (i) At step r, there are $m-r$ unmatched in-stubs left. Each unmatched in-stub can form a self-loop by connecting to an unmatched out-stub of the same vertex. The number of unmatched out-stubs at each vertex is upper bounded by $d_{\max}$, hence $\Delta_r^1 \le (m-r)d_{\max}$. The vertex to which an unmatched in-stub belongs has at most $d_{\max}-1$ incoming edges. The source of such an edge has at most $d_{\max}-1$ unmatched out-stubs left. Thus the number of out-stubs an unmatched in-stub can be paired with to create a double edge is at most $(d_{\max}-1)^2$. Hence $\Delta_r^2 \le (m-r)(d_{\max}-1)^2$ and $\Delta_r = \Delta_r^1 + \Delta_r^2 \le (m-r)d_{\max}^2$.

(ii) By definition, $\Lambda_r^{1+} = \sum_{i=1}^n d_i^{+(r)} d_i^+$. As $\sum_{i=1}^n d_i^{+(r)} = m-r$ and $d_i^+ \le d_{\max}$ for all i, this implies that $\Lambda_r^{1+} \le d_{\max}(m-r)$ and $\Lambda_r^{1-} \le d_{\max}(m-r)$.

(iii) By definition, $\Lambda_r = \sum_{(u,v)\in E_r} d_u^{+(r)} d_v^{-(r)} \frac{d_u^+ d_v^-}{2m} \le \frac{d_{\max}^2}{2m} \sum_{(u,v)\in E_r} d_u^{+(r)} d_v^{-(r)}$. Since $\sum_{i=1}^n d_u^{+(r)} = m-r$ and $d_v^{-(r)} \le (m-r)$ for all v, the claim follows. □

We will now determine the expected values of $\Delta_r^1, \Delta_r^2, \Lambda_r^{1+}, \Lambda_r^{1-}, \Lambda_r^2$ and Λ_r^3 in the G_{p_r} model, and, by combining this values together, will write the expression for ψ_r.

Lemma 2.4. *For each* $0 \le r \le m-1$ *the following equations hold:*

(i) $\mathbb{E}_{p_r}\left[\Delta_r^1\right] = \frac{(m-r)^2}{m^2} \sum_{i=1}^n d_i^+ d_i^-$*;*

(ii) $\mathbb{E}_{p_r}\left[\Delta_r^2\right] = \frac{r(m-r)^2}{m^3} \sum_{(i,j)\in G_{\mathbf{d}}} (d_i^+ - 1)(d_j^- - 1)$*;*

(iii) $\mathbb{E}_{p_r}\left[\Lambda_r^{1-}\Lambda_r^{1+}\right] = \frac{(m-r)^2}{m^2} \sum_{i=1}^n (d_i^-)^2 \sum_{i=1}^n (d_i^+)^2 + \frac{r(m-r)}{m^2} \sum_{(i,j)\in G_{\mathbf{d}}} d_i^+ d_j^-$*;*

(iv) $\mathbb{E}_{p_r}\left[\Lambda_r^2\right] = \frac{(m-r)^2}{m^2} \sum_{i=1}^n (d_i^-)^2 (d_i^+)^2$*;*

(v) $\mathbb{E}_{p_r}\left[\Lambda_r^3\right] = \frac{r(m-r)^2}{m^3} \sum_{(i,j)\in G_{\mathbf{d}}} d_i^+ (d_i^+ - 1) d_j^- (d_j^- - 1)$.

Proof. (i) The value of $d_i^{+(r)}$ equals the number of edges $(i,\bullet) \in G_{\mathbf{d}}$, such that $(i,\bullet) \notin G_{p_r}$. Since $p_r = \frac{r}{m}$, we have $\mathbb{E}_{p_r}\left[d_i^{\pm(r)}\right] = d_i^\pm \frac{m-r}{m}$. Furthermore, since $G_{\mathbf{d}}$ is simple, it contains no self-loops. This implies that $d_i^{-(r)}$ and $d_i^{+(r)}$ are independent. Using the fact that $\Delta_r^1 = \sum_{i=1}^n d_i^{-(r)} d_i^{+(r)}$, we find $\mathbb{E}_{p_r}\left[\Delta_r^1\right] = \frac{(m-r)^2}{m^2} \sum_{i=1}^n d_i^+ d_i^-$.

(ii) Δ_r^2 counts the number of pairs leading to a double edge. Choose a random $(i,j) \in G_{\mathbf{d}}$. To add an additional copy of this edge at step r, the edge must be already present in G_{p_r}, which happens with probability p_r. Let a pair of edges $(i,k),(l,j)$ be in $G_{\mathbf{d}}$ but not in G_{p_r}. This means that in G_{p_r} there are unmatched in-stubs and out-stubs such that one could instead form the edges (i,j) and (l,k), creating a double edge. The number of combinations of such l and k, is $(d_i^{+(r)}-1)(d_j^{-(r)}-1)$. By taking the expected value of this value, summing it over all edges of $G_{\mathbf{d}}$ and multiplying it by the probability p_r that $(i,j) \in G_{p_r}$, the claimed expected value of Δ_r^2 follows.

(iii) Remark that $\Lambda_r^{1-}\Lambda_r^{1+} = \sum_{j=1}^n \sum_{i=1}^n d_i^{+(r)} d_j^{-(r)} d_i^+ d_j^-$, which implies that

$$\mathbb{E}_{p_r}\left[\Lambda_r^{1-}\Lambda_r^{1+}\right] = \sum_{j=1}^n \sum_{i=1}^n \mathbb{E}_{p_r}\left[d_i^{+(r)} d_j^{-(r)}\right] d_i^+ d_j^- .$$

The random variables $d_i^{+(r)}$ and $d_j^{-(r)}$ are independent, unless $(i,j) \in G_{\mathbf{d}}$. Indeed, $d_i^{+(r)}$ (respectively $d_j^{-(r)}$) is the sum of d_i^+ (d_j^-) independent Bernoulli variables representing the out-stubs (in-stubs). If $(i,j) \in G_{\mathbf{d}}$, one fixed in-stub of j forms an edge with a fixed out-stub of i. This implies that the corresponding Bernoulli variables always need to take on the same value. Let us denote these Bernoulli variables by $d_{i_j}^+$ and $d_{j_i}^-$. Now that we have characterised the dependence between $d_i^{+(r)}$ and $d_j^{-(r)}$, we are ready to determine $\mathbb{E}_{p_r}\left[d_i^{+(r)} d_j^{-(r)}\right] = \mathbb{E}_{p_r}\left[d_i^{+(r)}\right] \mathbb{E}_{p_r}\left[d_j^{-(r)}\right] + \mathrm{Cov}\left(d_i^{+(r)} d_j^{-(r)}\right)$. As already explained in (i) $\mathbb{E}_{p_r}\left[d_i^{+(r)}\right] \mathbb{E}\left[d_j^{-(r)}\right] = \frac{(m-r)^2}{m^2} d_i^+ d_j^-$. For the covariance there holds

$$\mathrm{Cov}\left(d_i^{+(r)} d_j^{-(r)}\right) = \begin{cases} 0 & \text{if } (i,j) \notin G_{\mathbf{d}} \\ \mathrm{Cov}\left(d_{i_j}^+ d_{j_i}^-\right) & \text{if } (i,j) \in G_{\mathbf{d}} \end{cases} .$$

The covariance of any random variable X and a Bernoulli variable Y with expectation p^* equals: $\mathrm{Cov}\,(X,Y) = (\mathbb{E}\,[X|Y=1] - \mathbb{E}\,[X|Y=0])\, p^*(1-p^*)$. Applying this to $X = d_{i_j}^+$ and $Y = d_{j_i}^-$, their covariance becomes $\frac{r(m-r)}{m^2}$. Thus there holds

$$\mathbb{E}_{p_r}\left[d_i^{+(r)} d_j^{-(r)}\right] = \begin{cases} \frac{(m-r)^2}{m^2} d_i^+ d_j^- & \text{if } (i,j) \notin G_{\mathbf{d}} \\ \frac{(m-r)^2}{m^2} d_i^+ d_j^- + \frac{r(m-r)}{m^2} & \text{if } (i,j) \in G_{\mathbf{d}} \end{cases} .$$

Plugging this back into the expression for $\mathbb{E}_{p_r}\left[\Lambda_r^{1-}\Lambda_r^{1+}\right]$ the desired equation follows.

(iv) Recall that $\Lambda_r^2 = \sum_{i=1}^n d_i^{-(r)} d_i^{+(r)} d_i^- d_i^+$. In the proof of (i) we have already showed that $\mathbb{E}_{p_r}\left[d_i^{-(r)} d_i^{+(r)}\right] = d_i^+ d_i^- \frac{(m-r)^2}{m^2}$. Hence $\mathbb{E}_{p_r}\left[\Lambda_r^2\right] = \frac{(m-r)^2}{m^2} \sum_{i=1}^n {d_i^-}^2 {d_i^+}^2$.

(v) From equation (12) it follows that $\Lambda_r^3 = \sum_{(i,j) \notin E_r, i \neq j} d_i^+ d_j^- d_i^{+(r)} d_j^{-(r)}$.
Since that $\Delta_r^2 = \sum_{(i,j) \notin E_r, i \neq j} d_i^{+(r)} d_j^{-(r)}$ we can use the proof of (ii). This implies each edge $(i,j) \in G_{\mathbf{d}}$ contributes $\frac{(m-r)^2}{m^2} \frac{r d_i^+ (d_i^+ - 1) d_j^- (d_j^- - 1)}{m}$ to the sum, proving the claim.

□

Next, we will use the following asymptotic estimates,

a) $\sum_{i=1}^{n} \left(d_i^-\right)^s = \sum_{(i,j)\in G_{\mathbf{d}}} \left(d_i^-\right)^{s-1} = \mathcal{O}\left(m d_{\max}^{s-1}\right)$,

b) $\sum_{i=1}^{n} \left(d_i^+\right)^t = \sum_{(i,j)\in G_{\mathbf{d}}} \left(d_i^+\right)^{t-1} = \mathcal{O}\left(m d_{\max}^{t-1}\right)$,

c) $\sum_{i=1}^{n} \left(d_i^-\right)^s \left(d_i^-\right)^t = \sum_{(i,j)\in G_{\mathbf{d}}} \left(d_i^-\right)^{s-1} \left(d_i^+\right)^t = \mathcal{O}\left(m d_{\max}^{s+t-1}\right)$,

to obtain and approximation of ψ_r that we will work with. Combing these estimates with Lemma 2.4 we find

$$\mathbb{E}_{p_r}\left[\frac{\Lambda_r^{1-}\Lambda_r^{1+}}{4m}\right] = \frac{(m-r)^2}{4m^3}\sum_{i=1}^{n}(d_i^-)^2\sum_{i=1}^{n}(d_i^+)^2 + (m-r)^2\mathcal{O}\left(\frac{r d_{\max}^2}{(m-r)m^2}\right),$$

$$\mathbb{E}_{p_r}\left[\frac{\Lambda_r^2}{4m}\right] = (m-r)^2\mathcal{O}\left(\frac{d_{\max}^3}{m^2}\right) \quad \text{and} \quad \mathbb{E}_{p_r}\left[\frac{\Lambda_r^3}{2m}\right] = (m-r)^2\mathcal{O}\left(r\frac{d_{\max}^4}{m^3}\right).$$

This allows us to state the following Lemmas, which will be useful in Sections 2.3 and 2.4.

Lemma 2.5. *For all $0 \le r \le m-1$ there holds:*

$$\psi_r = (m-r)^2\left[\frac{\sum_{i=1}^{n} d_i^- d_i^+}{m^2} + \frac{r\sum_{(i,j)\in G_{\mathbf{d}}}\left(d_i^+ - 1\right)\left(d_j^- - 1\right)}{m^3} + \frac{\sum_{i=1}^{n}(d_i^-)^2\sum_{i=1}^{n}(d_i^+)^2}{4m^3} + \xi_r\right], \tag{13}$$

with error term $\xi_r = \mathcal{O}\left(\frac{d_{\max}^3}{m^2} + \frac{r d_{\max}^2}{(m-r)m^2} + \frac{r d_{\max}^4}{m^3}\right)$.

Lemma 2.6. *For each $0 \le r \le m-1$ the quantity ψ_r is upper bounded by $\mathcal{O}\left((m-r)^2\frac{d_{\max}^2}{m}\right)$.*

Proof. Combing equation (13) with the asymptotic estimate

$$\sum_{i=1}^{n}\left(d_i^-\right)^s\left(d_i^-\right)^t = \sum_{(i,j)\in G_{\mathbf{d}}}\left(d_i^-\right)^{s-1}\left(d_i^+\right)^t = \mathcal{O}\left(m d_{\max}^{s+t-1}\right)$$

we find that $\psi_r = (m-r)^2\mathcal{O}\left(\frac{d_{\max}}{m} + \frac{r d_{\max}^2}{m^2} + \frac{d_{\max}^2}{4m} + \frac{r d_{\max}^2}{(m-r)m^2} + \frac{d_{\max}^3}{m^3} + \frac{r d_{\max}^4}{m^3}\right)$, and since $r \le m$ and $d_{\max}^2 = o(m)$, the latter equation becomes

$$\psi_r = (m-r)^2\mathcal{O}\left(\frac{d_{\max}^2}{m}\right).$$

□

2.3 Proving equation (5)

With help of Lemmas 2.5 and 2.6 we are now ready to prove equation (5). We start by multiplying the left hand side of equation (5) by $\prod_{r=0}^{m-1}(m-r)^2$. This leads to

$$\prod_{r=0}^{m-1}\frac{(m-r)^2}{(m-r)^2-\psi_r}=\prod_{r=0}^{m-1}\left(1+\frac{\psi_r}{(m-r)^2-\psi_r}\right).$$

Applying Lemma 2.5 to the numerator and Lemma 2.6 to the denominator we the right hand side of the later equation becomes:

$$\exp\left[\sum_{r=0}^{m-1}\ln\left(1+\frac{\frac{\sum_{i=1}^n d_i^- d_i^+}{m^2}+\frac{r\sum_{(i,j)\in G_{\mathbf{d}}}\left(d_i^+-1\right)\left(d_j^--1\right)}{m^3}+\frac{\sum_{i=1}^n(d_i^-)^2\sum_{i=1}^n(d_i^+)^2}{4m^3}+\xi_r}{1-\mathcal{O}\left(\frac{d_{\max}^2}{m}\right)}\right)\right].$$

and after using that $\mathcal{O}\left(\frac{d_{\max}^2}{m}\right)=\mathcal{O}\left(\frac{1}{m^{1/2+2\tau}}\right)$ and some asymptotic expansions, we obtain:

$$\prod_{r=0}^{m-1}\frac{(m-r)^2}{(m-r)^2-\psi_r}=[1+o(1)]\exp\left[\frac{\sum_{i=1}^n d_i^- d_i^+}{m}-\frac{\sum_{i=1}^n(d_i^-)^2+\sum_{i=1}^n(d_i^+)^2}{2m}+\right.$$
$$\left.\frac{\sum_{i=1}^n(d_i^-)^2\sum_{i=1}^n(d_i^+)^2}{4m^2}+\frac{\sum_{(i,j)\in G_{\mathbf{d}}}d_i^+d_j^-}{2m}+\frac{1}{2}\right],$$

which proves equation (5).

2.4 Proving equation (4)

Let us define

$$f(\mathcal{N}):=\prod_{r=0}^{m-1}\frac{(m-r)^2-\psi_r}{(m-r)^2-\Psi_r}. \tag{14}$$

Then equation (4) becomes equivalent to

$$\mathbb{E}\left[f(\mathcal{N})\right]=1+o(1), \tag{15}$$

which we will demonstrate instead in the remainder of this section. We start by rewriting the latter expectation as a sum of expected values

$$\mathbb{E}\left[f(\mathcal{N})\right]=\mathbb{E}\left[f(\mathcal{N})\,\mathbb{1}_{\mathcal{A}}\right]+\mathbb{E}\left[f(\mathcal{N})\,\mathbb{1}_{\mathcal{B}}\right]+\mathbb{E}\left[f(\mathcal{N})\,\mathbb{1}_{\mathcal{C}}\right]+\mathbb{E}\left[f(\mathcal{N})\,\mathbb{1}_{S(\mathcal{M})\setminus S^*(\mathcal{M})}\right],$$

of mutually disjoint subsets covering $S(\mathcal{M})$ in the following fashion.

Partitioning $S(\mathcal{M})$ The set of orderings $S(\mathcal{M})$ is partitioned as follows:

1. For a small number $0\le\tau\le\frac{1}{3}$, such that $d_{\max}=\mathcal{O}\left(m^{1/4-\tau}\right)$, we define

$$S^*(\mathcal{M})=\left\{\mathcal{N}\in S(\mathcal{M})\,|\,\Psi_r(\mathcal{N})-\psi_r\le\left(1-\frac{\tau}{4}\right)(m-r)^2,\forall\,0\le r\le m-1\right\}, \tag{16}$$

and let $S(\mathcal{M})\setminus S^*(\mathcal{M})$ be the first element of the partition.

2. As the second element of the partition we take

$$\mathcal{A} = \left\{\mathcal{N} \in S^*(\mathcal{M}) \,|\, \Psi_r(\mathcal{N}) - \psi_r > T_r\left(\ln(n)^{1+\delta}\right), \forall\, 0 \le r \le m-1\right\}, \tag{17}$$

where the family of functions T_r is defined bellow, see equation (25), and δ is a small positive constant, *e.g.* $0 < \delta < 0.1$.

3. The next element of the partition is chosen from $S^*(\mathcal{M}) \setminus \mathcal{A}$ to be

$$\mathcal{B} = \left\{\mathcal{N} \in S^*(\mathcal{M}) \setminus \mathcal{A} \,|\, \exists 0 \le r \le m-1, \text{ s.t. } m - r \le \ln(n)^{1+2\delta} \text{ and } \Psi_r(\mathcal{N}) > 1\right\}. \tag{18}$$

4. We define as last element as the complement

$$\mathcal{C} = S^*(\mathcal{M}) \setminus (\mathcal{A} \cup \mathcal{B}). \tag{19}$$

We will now show that the following asymptotic estimates hold

$$\mathbb{E}\left(f(\mathcal{N})\, \mathbb{1}_{\mathcal{A}}\right) = o(1); \tag{20}$$
$$\mathbb{E}\left(f(\mathcal{N})\, \mathbb{1}_{\mathcal{B}}\right) = o(1); \tag{21}$$
$$\mathbb{E}\left(f(\mathcal{N})\, \mathbb{1}_{\mathcal{C}}\right) \le 1 + o(1); \tag{22}$$
$$\mathbb{E}\left(f(\mathcal{N})\, \mathbb{1}_{\mathcal{C}}\right) \ge 1 - o(1); \tag{23}$$
$$\mathbb{E}\left(f(\mathcal{N})\, \mathbb{1}_{S(\mathcal{M})\setminus S^*(\mathcal{M})}\right) = o(1). \tag{24}$$

Since $\mathbb{E}\left[f(\mathcal{N})\right]$ is as sum of the above expected values, it remains to introduce suitable definitions for T_r and prove equations (20)-(24) to finish the proof of (15).

The family of functions T_r. We define the family of functions $T_r : \mathbb{R}_{\ge 0} \to \mathbb{R}_{\ge 0}$ indexed by $r \in \{0, 1, \ldots, m-1\}$ as follows

$$\mathrm{T}_r(\lambda) := \begin{cases} 4\beta_r(\lambda) + 2\min(\gamma_r(\lambda), \nu_r) & \text{if } m - r \ge \lambda\omega, \\ \frac{\lambda^2}{\omega^2}, & \text{otherwise,} \end{cases} \tag{25}$$

with

$$\beta_r(\lambda) := c\sqrt{\lambda\left(m d_{\max}^2 q_r^2 + \lambda^2\right)\left(d_{\max}^2 q_r + \lambda\right)}, \tag{26}$$
$$\gamma_r(\lambda) := c\sqrt{\lambda\left(m d_{\max}^2 q_r^3 + \lambda^3\right)\left(d_{\max}^2 q_r^2 + \lambda^2\right)}, \tag{27}$$
$$\nu_r := 8 m d_{\max}^2 q_r^3, \tag{28}$$
$$\omega := \ln(n)^{\delta}, \tag{29}$$
$$q_r := \frac{m-r}{m} = 1 - p_r. \tag{30}$$

The quantity c is a large positive constant, which will be defined later, and q_r is the probability that an edge of $G_{\mathbf{d}}$ is not present in G_{p_r}. The intuition behind the definition of this family of functions will become apparent in the remainder of this section. Let $\lambda_0 := \omega \ln(n)$ and $\lambda_i := 2^i \lambda_0$ for all $i \in \{1, 2, \ldots, L\}$, where L is the unique integer such that $\lambda_{L-1} < c\, d_{\max} \ln(n) \le \lambda_L$. There holds the following relation between $T_r(\lambda_i)$ and $T_r(\lambda_{i-1})$.

Lemma 2.7. *For all* $0 \le r \le m-1$ *and* $i \in \{1, 2, \dots, L\}$ *there holds*

$$T_r(\lambda_i) \le 8 T_r(\lambda_{i-1}).$$

Proof. As the function T_r is defined piecewise, we distinguish three cases:

1. Suppose $m - r < \lambda_i \omega$ and $m - r < \lambda_{i-1}\omega$.
   Then
   $$T_r(\lambda_i) = \frac{\lambda_i^2}{\omega^2} = \frac{4\lambda_{i-1}^2}{\omega^2} < \frac{8\lambda_{i-1}^2}{\omega^2} = 8T_r(\lambda_{i-1}),$$
   showing that $T_r(\lambda_i) \le 8T_r(\lambda_{i-1})$.

2. Suppose $m - r < \lambda_i \omega$ and $m - r \ge \lambda_{i-1}\omega$.
   Then by definition of T_r there holds $T_r(\lambda_i) = \frac{4\lambda_{i-1}^2}{\omega^2}$ and $T_r(\lambda_{i-1}) \ge 4\beta_r(\lambda_{i-1}) \ge 4c\lambda_{i-1}^2$. Hence we find $T_r(\lambda_i) \le T_r(\lambda_{i-1})$.

3. Suppose $m - r \ge \lambda_i \omega$ and $m - r \ge \lambda_{i-1}\omega$.
   Then by definition of T_r, there holds $T_r(\lambda_i) = 4\beta_r(\lambda_i) + 2\min(\gamma_r(\lambda_i), \nu_r)$ and $T_r(\lambda_{i-1}) = 4\beta_r(\lambda_{i-1}) + 2\min(\gamma_r(\lambda_{i-1}), \nu_r)$. Both $\beta_r(\lambda)$ and $\gamma_r(\lambda)$ are square roots of a 6^{th}-order polynomial in λ. As $\lambda_i = 2\lambda_{i-1}$ and $\sqrt{2^6} = 8$, this implies that $\beta_r(\lambda_i) \le 8\beta_r(\lambda_{i-1})$ and $\gamma_r(\lambda_i) \le 8\gamma_r(\lambda_{i-1})$. Hence there holds $T_r(\lambda_i) \le 8T_r(\lambda_{i-1})$.

This completes the proof, because $m - r \ge \lambda_i \omega$ and $m - r < \lambda_{i-1}\omega$ never holds, as $\lambda_i > \lambda_{i-1}$. □

In order to prove equations (20) and (21) we subpartition $\mathcal{A}$ and $\mathcal{B}$. Let us define the chain of subsets $A_0 \subset A_1 \subset \ldots \subset A_L \subset S^*(\mathcal{M})$ with

$$A_i = \{\mathcal{N} \in S^*(\mathcal{M}) \mid \Psi_r(\mathcal{N}) - \psi_r < T_r(\lambda_i), \forall 0 \le r \le m-1\}. \tag{31}$$

To ensure that we cover $S^*(\mathcal{M})$ entirely, we also introduce

$$A_\infty = S^*(\mathcal{M}) \setminus A_L = \{\mathcal{N} \in S^*(\mathcal{M}) \mid \exists\, 0 \le r \le m-1, \text{ s.t. } \Psi_r(\mathcal{N}) - \psi_r \ge T_r(\lambda_L)\}. \tag{32}$$

Now equation (17) implies that

$$\mathcal{A} = S^*(\mathcal{M}) \setminus A_0 = \cup_{i=1}^L A_i \setminus A_{i-1} \bigcup A_\infty.$$

Next, we partition A_0. The goal of this partition is to write $\mathcal{B}$ as the union of some smaller sets. As $\mathcal{N} \in A_0$ for all $0 \le r \le m-1$ such that $r \ge m - \omega\lambda_0$ there holds

$$\Psi_r(\mathcal{N}) < T_r(\lambda_0) + \psi_r = \ln(n)^2 + \psi_r.$$

According to Lemma 2.6 for all $m - 1 \ge r \ge m - \omega\lambda_0$, $\psi_r = o(1)$. Hence there exists some n_0 such that for all $n > n_0$:

$$\Psi_r(\mathcal{N}) < \ln(n)^2 + 1.$$

Without loss of generality we may assume that $n > n_0$. Let K be the unique integer such that

$$2^{K-1} < \ln(n)^2 + 1 \le 2^K. \tag{33}$$

Then for all $r \geq m - \omega\lambda_0$:

$$\Psi_r \leq 2^K.$$

This allows us to define the chain of subsets $B_0 \subset B_1 \subset \ldots \subset B_K = A_0$, with

$$B_j = \left\{\mathcal{N} \in A_0 | \, \Psi_r(\mathcal{N}) < 2^j, \forall r \geq m - \omega\lambda_0\right\}. \tag{34}$$

From equations (18) and (19) it immediately follows that

$$\mathcal{B} = \cup_{i=1}^{K} B_i \setminus B_{i-1} \quad \text{and} \quad \mathcal{C} = B_0.$$

These descriptions of $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ enable us to show validity of equations (20), (21), (22) and (23). First, we prove equation (20). The proof also contains statements that hold for any ordering in $S^*(\mathcal{M})$, which are also used in the proof of equations (21), (22) and (23). We finish with the proof of equation (24), which requires a different technique as it concerns all orderings not in $S^*(\mathcal{M})$.

Proving equation (20) Based on the definition of $\mathcal{A}$ in terms of A_i's and A_∞, we now prove equation (20). For this we use the following Lemmas.

Lemma 2.8. *For all $1 \leq i \leq L$ there holds*

(a) $\mathbb{P}[\mathcal{N} \in A_i \setminus A_{i-1}] \leq e^{-\Omega(\lambda_i)}$*;*

(b) For all $\mathcal{N} \in A_i \setminus A_{i-1}$ there holds $f(\mathcal{N}) \leq e^{o(\lambda_i)}$.

Lemma 2.9. *For a large enough constant c there holds*

(a) $\mathbb{P}[\mathcal{N} \in A_\infty] \leq e^{-\Omega(cd_{\max}\ln(n))}$*;*

(b) For all $\mathcal{N} \in A_\infty$ there holds $f(\mathcal{N}) \leq e^{72d_{\max}\ln(n)}$.

Together these lemmas imply that

$$\mathbb{E}[f(\mathcal{N})\mathbb{1}_{\mathcal{A}}] \leq \sum_{i=1}^{L} e^{-\Omega(\lambda_i)}e^{o(\lambda_i)} + e^{-\Omega(cd_{\max}\ln(n))}e^{72d_{\max}\ln(n)} = o(1),$$

thus proving equation (20).

First, we prove Lemma 2.8 (a) and Lemma 2.9 (a). This is done by showing a stronger statement,

$$\mathbb{P}\left[\mathcal{N} \in A_{i-1}^c\right] \leq e^{-\Omega(\lambda_i)},$$

for all $i \in \{0, 1, \ldots, L\}$. This statement is indeed stronger than the statements of Lemma 2.8 (a) as $(A_i \setminus A_{i-1}) \subset (S(\mathcal{M}) \setminus A_{i-1})$. This observation is also relevant for Lemma 2.9 (a) since $A_\infty \in A_L^c$ and $\lambda_L \geq cd_{\max}\ln(n)$. Combining the definition of A_{i-1} with Lemma 2.7, we find

$$A_{i-1}^c \subset \left\{\mathcal{N} \in S(\mathcal{M}) \,|\, \exists\, 0 \leq r \leq m-1 \,\text{s.t.}\, \Psi_r(\mathcal{N}) - \psi_r > \frac{T_r(\lambda_i)}{8}\right\}.$$

This implies that to prove Lemma 2.8 (a) and Lemma 2.9 (a), it suffices to show that for all $i \in \{0, 1, \dots, L\}$ and $0 \le r \le m-1$,

$$\mathbb{P}\left[|\Psi_r - \psi_r| \ge \frac{T_r(\lambda_i)}{8}\right] \le e^{-\Omega(\lambda_i)}. \tag{35}$$

Determining the value of Ψ_r is more challenging than the value of Ψ_{p_r} in random graph model G_{p_r}, where each edge is present with probability p_r. As mentioned in Section 2.2, the graph $G_{\mathcal{N}_r}$ is a random subgraph of $G_{\mathbf{d}}$ with exactly r edges for a random ordering $\mathcal{N} \in S(\mathcal{M})$. Denoting the number of edges in G_{p_r} by $E[G_{p_r}]$ we find:

$$\mathbb{P}\left[|\Psi_r - \psi_r| \ge \frac{T_r(\lambda_i)}{8}\right] = \frac{\mathbb{P}\left[|\Psi_{p_r} - \psi_r| \ge \frac{T_r(\lambda_i)}{8} \cap |E[G_{p_r}]| = r\right]}{\mathbb{P}\left[|E[G_{p_r}]| = r\right]} \le \frac{\mathbb{P}\left[|\Psi_{p_r} - \psi_r| \ge \frac{T_r(\lambda_i)}{8}\right]}{\mathbb{P}\left[|E[G_{p_r}]| = r\right]}.$$

Bayati, Kim and Saberi showed the following bound on the probability that the random graph G_{p_r} contains exactly r edges.

Lemma 2.10. *([15, Lemma 21]) For all $0 \le r \le m$ there holds $\mathbb{P}\left[|E[G_{p_r}]| = r\right] \ge \frac{1}{n}$.*

Using this Lemma we obtain

$$\mathbb{P}\left[|\Psi_r - \psi_r| \ge \frac{T_r(\lambda_i)}{8}\right] \le n \cdot \mathbb{P}\left[|\Psi_{p_r} - \psi_r| \ge \frac{T_r(\lambda_i)}{8}\right].$$

As $\lambda_i = 2^i \ln(n)^{1+\delta} \gg \ln(n)$, there holds $ne^{-\Omega(\lambda_i)} = e^{-\Omega(\lambda_i)+\ln(n)} = e^{-\Omega(\lambda_i)}$. Hence, to prove equation (35) it suffices to show that

$$\mathbb{P}\left[|\Psi_{p_r} - \psi_r| \ge \frac{T_r(\lambda_i)}{8}\right] \le e^{-\Omega(\lambda_i)}.$$

As T_r is defined piecewise, we formulate separate Lemmas distinguishing two cases: *i)* $m - r < \omega\lambda_i$ and *ii)* $m - r \ge \omega\lambda_i$.

Lemma 2.11. *For all $i \in \{0, 1, \dots, L\}$ and $0 \le r \le m-1$ such that $m - r < \lambda_i\omega$ there holds*

$$\mathbb{P}\left[\Psi_{p_r} - \psi_r \ge \frac{\lambda_i^2}{8\omega^2}\right] \le e^{-\Omega(\lambda_i)}. \tag{36}$$

Proof. Instead of showing the desired inequality, we show an even stronger statement:

$$\mathbb{P}\left[\Psi_{p_r} \ge \frac{\lambda_i^2}{8\omega^2}\right] \le e^{-\Omega(\lambda_i)}.$$

Combining the fact that $\Psi_{p_r} \le \frac{\lambda_i^2}{8\omega}$ with $\Psi_{p_r} = \Delta_{p_r} + \Lambda_{p_r}$ and Lemma 2.3, we find

$$\Delta_{p_r} \ge \frac{\lambda_i^2}{8\omega^2} - \frac{d_{\max}^2 m}{2} q_r^2.$$

As $mq_r = m - r < \omega\lambda_i$ and $\omega^4 d_{\max}^2 < \frac{m}{5}$ for large n there holds

$$\Delta_{p_r} \ge \frac{\lambda_i^2}{8\omega^2} - \frac{d_{\max}^2}{2m}\omega^2\lambda_i^2 \quad \ge \frac{\lambda_i^2}{40\omega^2}.$$

Let G_{q_r} be the complement of G_{p_r} in $G_{\mathbf{d}}$ and define $N_0(u) := \{v \in V \mid (u,v) \in G_{q_r}\} \cup \{u\}$. Let $d^+_{G_{q_r}}(u)$ (respectively $d^-_{G_{q_r}}(u)$) be the out-degree (in-degree) of u in G_{q_r}. By definition of Δ_{p_r} there holds

$$\Delta_{p_r} \leq \sum_{u \in V} d^+_{G_{q_r}}(u) \sum_{v \in N_0(u)} d^-_{G_{q_r}}(v).$$

By combining the latter inequality with the lower bound on Δ_{p_r} we have just derived, we find

$$\frac{\lambda_i^2}{40\omega^2} \leq \Delta_{p_r} \leq \sum_{u \in V} d^+_{G_q}(u) \sum_{v \in N_0(u)} d^-_{G_q}(v). \tag{37}$$

This equation implies that at least one of the following statements must hold true:

(a) G_q has more than $\frac{\omega^2 \lambda_i}{40}$ edges;

(b) for some $u \in V$ there holds $\sum_{v \in N_0(u)} d^-_{G_q}(v) \geq \frac{\lambda_i}{\omega^4}$.

If (a) is violated, there holds $\sum_{u \in V} d^+_{G_q}(u) \leq \frac{\omega^2 \lambda_i}{40}$. If (b) is violated, for all $u \in V$ there holds $\sum_{v \in N_0(u)} d^-_{G_q}(v) < \frac{\lambda_i}{\omega^4}$. Hence if (a) and (b) are both violated, we find

$$\Delta_{p_r} \leq \sum_{u \in V} d^+_{G_q}(u) \sum_{v \in N_0(u)} d^-_{G_q}(v) < \frac{\omega^2 \lambda_i}{40} \frac{\lambda_i}{\omega^4} = \frac{\lambda_i^2}{40\omega^2}.$$

This violates equation (37). Thus it is not possible that (a) and (b) are simultaneously violated. This implies that at least one of the statements holds. Using the proof of [15, Lemma 20], the probabilities that statements (a) and (b) hold, are both upper bounded by $e^{-\Omega(\lambda_i)}$. Since $\Psi_{p_r} \geq \frac{\lambda_i^2}{8\omega}$ implies that at least one of these statements holds, this completes the proof. □

Lemma 2.12. *For all $i \in \{0, 1, \dots, L\}$ and r such that $m - r \geq \lambda_i \omega$ there holds*

$$\mathbb{P}\left[|\Psi_{p_r} - \psi_r| \geq \frac{4\beta_r(\lambda_i) + 2\min(\nu_r, \gamma_r(\lambda_i))}{8}\right] \leq e^{-\Omega(\lambda_i)}. \tag{38}$$

Recall that $\Psi_{p_r} = \Delta^1_{p_r} + \Delta^2_{p_r} + \frac{\Lambda^1_{p_r}{}^+ \Lambda^1_{p_r}{}^- - \Lambda^2_{p_r}}{4m} - \frac{\Lambda^3_{p_r}}{2m}$ and that ψ_r equals $\mathbb{E}[\Psi_{p_r}]$. Thus to prove Lemma 2.12, it suffices to concentrate $\Delta^1_{p_r}, \Delta^2_{p_r}, \Lambda^1_{p_r}{}^+ \Lambda^1_{p_r}{}^-, \Lambda^2_{p_r}$ and $\Lambda^3_{p_r}$ around their expected values with probability $e^{-\Omega(\lambda_i)}$ such that the difference between their sum and the sum of their expected values is smaller than $\frac{4\beta_r(\lambda_i)+2\min(\nu_r,\gamma_r(\lambda_i))}{8}$. This is shown using Vu's concentration inequality.

Theorem 2.13. *[Vu's concentration inequality [36]] Consider independent random variables $t_1, t_2, \dots, t_n$ with arbitrary distribution in $[0,1]$. Let $Y(t_1, t_2, \dots, t_n)$ be a polynomial of degree k with coefficients in $(0,1]$. For any multi-set A let $\partial_A Y$ denote the partial derivative with respect to the variables in A. Define $\mathbb{E}_j(Y) = \max_{|A| \geq j} \mathbb{E}(\partial_A Y)$ for all $0 \leq j \leq k$. Recursively define $c_1 = 1, d_1 = 2, c_k = 2\sqrt{k}(c_{k-1}+1), d_k = 2(d_{k-1}+1)$. Then for any $\mathcal{E}_0 > \mathcal{E}_1 > \ldots > \mathcal{E}_k = 1$ and λ fulfilling* *i)* $\mathcal{E}_j \geq \mathbb{E}_j(Y)$*;*

ii) $\frac{\mathcal{E}_j}{\mathcal{E}_{j-1}} \geq \lambda + 4j\ln(n)$ *for all* $0 \leq j \leq k-1$*;*

there holds

$$\mathbb{P}\left[|Y - \mathbb{E}[Y]| \geq c_k\sqrt{\lambda\mathcal{E}_0\mathcal{E}_1}\right] \leq d_k e^{-\lambda/4}.$$

Lemma 2.14. *For all* $i \in \{0, 1, \dots, L\}$ *and* $0 \leq r \leq m-1$ *there holds:*

(i) $\mathbb{P}\left[\left|\Delta^1_{p_r} - \mathbb{E}\left[\Delta^1_{p_r}\right]\right| \geq \frac{\beta_r(\lambda_i)}{8}\right] \leq e^{-\Omega(\lambda_i)}$*;*

(ii) $\mathbb{P}\left[\left|\Delta^2_{p_r} - \mathbb{E}\left[\Delta^2_{p_r}\right]\right| \geq \frac{\min(\beta_r(\lambda_i)+\gamma_r(\lambda_i), \beta_r(\lambda_i)+\nu_r)}{8}\right] \leq e^{-\Omega(\lambda_i)}$*;*

(iii) $\mathbb{P}\left[\left|\frac{\Lambda^1_{p_r}{}^- \Lambda^1_{p_r}{}^+ - \Lambda^2_{p_r}}{4m} - \frac{\mathbb{E}\left[\Lambda^1_{p_r}{}^- \Lambda^1_{p_r}{}^+ - \Lambda^2_{p_r}\right]}{4m}\right| \geq \frac{\beta_r(\lambda_i)}{8}\right] \leq e^{-\Omega(\lambda_i)}$*;*

(iv) $\mathbb{P}\left[\left|\frac{\Lambda^3_{p_r}}{2m} - \frac{\mathbb{E}\left[\Lambda^3_{p_r}\right]}{2m}\right| \geq \frac{\min(\beta_r(\lambda_i)+\gamma_r(\lambda_i), \beta_r(\lambda_i)+\nu_r)}{8}\right] \leq e^{-\Omega(\lambda_i)}$*.*

Proof. To prove each of the above equations, we write the quantity as a polynomial and apply Theorem 2.13 to it. This polynomial will be a function of m Bernoulli variables. Each variable t_e represents an edge $e \in G_{\mathbf{d}}$, that is if $e \in G_{p_r}$ then $t_e = 0$ and if $e \notin G_{p_r}$, $t_e = 1$. Remark that by definition of G_{p_r}, see Section 2.2, there holds $\mathbb{E}[t_e] = q_r$ for all e. Also by definition of G_{p_r}, the variables t_e are independent of each other.

(i) Recall that $\Delta^1_{p_r}$ counts the number of pairs creating a self-loop. Each vertex v has d^-_v in-stubs and d^+_v out-stubs. The number of those out-stubs (respectively in-stubs) that are matched equals the number of outgoing (incoming edges) for v in G_{p_r}. Thus the number of unmatched in-stubs (respectively out-stubs) of vertex v is $\sum_{e=(\bullet,v)\in G_{\mathbf{d}}} t_e$ $\left(\sum_{e=(v,\bullet)\in G_{\mathbf{d}}} t_e\right)$. The number of ways to create a self-loop at v is

$$\sum_{e=(v,\bullet)\in G_{\mathbf{d}}} \sum_{f=(\bullet,v)\in G_{\mathbf{d}}} t_e t_f.$$

Hence we find

$$\Delta^1_{p_r} = \sum_{v\in V} \sum_{e=(v,\bullet)\in G_{\mathbf{d}}} \sum_{f=(\bullet,v)\in G_{\mathbf{d}}} t_e t_f. \tag{39}$$

Vu's concentration inequality requires us to upper bound the values $\mathbb{E}_0\left[\Delta^1_{p_r}\right], \mathbb{E}_1\left[\Delta^1_{p_r}\right]$ and $\mathbb{E}_2\left[\Delta^1_{p_r}\right]$. Let us first consider the expectation of $\Delta^1_{p_r}$. Because $G_{\mathbf{d}}$ is simple, for each element of the summation in equation (39) e does not equal f. Therefore $\mathbb{E}[t_e t_f] = q_r^2$. The summations over v and e in equation (39), can be replaced by one summation over all edges in $G_{\mathbf{d}}$. For each edge $e \in G_{\mathbf{d}}$, there are at most $d_{\max}$ edges in $G_{\mathbf{d}}$ with the source of e as target. Hence we find $\mathbb{E}\left[\Delta^1_{p_r}\right] \leq m d_{\max} q_r^2$. Let us take the partial derivative with respect to one variable t_e for some $e = (u, v)$, then we obtain $\sum_{f=(\bullet,u)\in G_{\mathbf{d}}} t_f + \sum_{f=(v,\bullet)\in G_{\mathbf{d}}} t_f$. This is upper bounded by $2d_{\max}q_r$. As $\Delta^1_{p_r}$ is a polynomial of degree 2 with all coefficients 1, it is clear that $\mathbb{E}\left[\partial_{t_e}\partial_{t_f}\Delta^1_{p_r}\right] \leq 1$ for all e, f. Thus we find

$$\mathbb{E}_0\left[\Delta^1_{p_r}\right] \leq \max\left(1, 2d_{\max}q_r, md_{\max}q_r^2\right), \ \mathbb{E}_1\left[\Delta^1_{p_r}\right] \leq \max\left(1, 2d_{\max}q_r\right), \text{ and } \mathbb{E}_2\left[\Delta^1_{p_r}\right] \leq 1.$$

The maximization follows from the definition of $\mathbb{E}_j(Y)$. Let us define,

$$\mathcal{E}_0 := 9\lambda_i^2 + 2md_{\max}q_r^2, \quad \mathcal{E}_1 := 9\lambda_i + 2d_{\max}q_r \quad \text{and} \quad \mathcal{E}_2 := 1.$$

We claim that together with $\lambda = \lambda_i$, they fulfil the conditions of Theorem 2.13. It is obvious that $\mathcal{E}_2 \geq \mathbb{E}_2\left[\Delta^1_{p_r}\right]$. Also $\mathcal{E}_1 \geq \mathbb{E}_1\left[\Delta^1_{p_r}\right]$ as $\lambda_i \geq 1$ for all $n \geq 3$. Furthermore $\mathcal{E}_0 \geq \mathbb{E}_0\left[\Delta^1_{p_r}\right]$ as $\lambda_i \geq 1$ and $mq_r = m - r$ implies that $2md_{\max}q_r^2 \geq 2d_{\max}q_r$. This shows the first condition of Theorem 2.13. For the second condition remark that $\lambda_i \geq \ln(n)$ and $\ln(m) \leq 2\ln(n)$ as $m \leq n^2$. This implies

$$\frac{\mathcal{E}_1}{\mathcal{E}_2} = \mathcal{E}_1 \geq \lambda_i + 4\ln(m).$$

Furthermore, there holds

$$\frac{\mathcal{E}_0}{\mathcal{E}_1} = \lambda_i \left(\frac{9\lambda_i + \frac{2d_{\max}mq_r^2}{\lambda_i}}{9 + \frac{2d_{\max}q_r}{\lambda_i}} \right) \geq \lambda_i,$$

showing that the second condition of Theorem 2.13 is fulfilled as well. Thus we may apply Vu's concentration inequality to obtain

$$\mathbb{P}\left[\left|\Delta^1_{p_r} - \mathbb{E}\left[\Delta^1_{p_r}\right]\right| \geq c_2\sqrt{\lambda_i\left(9\lambda_i + 2d_{\max}q_r\right)\left(9\lambda_i^2 + 2md_{\max}q_r^2\right)}\right] \leq e^{-\Omega(\lambda_i)}.$$

Since $\mathbb{P}\left[\left|\Delta^1_{p_r} - \mathbb{E}\left[\Delta^1_{p_r}\right]\right| \geq a\right] \leq \mathbb{P}\left[\left|\Delta^1_{p_r} - \mathbb{E}\left[\Delta^1_{p_r}\right]\right| \geq b\right]$ for $a > b$, choosing any $c > 8 \cdot 9c_2$ in equation (26) completes the proof.

(ii) Recall that $\Delta^2_{p_r}$ counts the number of pairs that create an edge already present in G_{p_r}, *i.e.* a double edge. Pairing an out-stub of u with an in-stub of v creates a double edge only if $(u, v) \in G_{p_r}$, i.e. if for $e = (u, v)$, $t_e = 1$. Recalling the expressions for the number of unmatched in-stubs and out-stubs at a vertex v from the proof of (i) and defining a set of non-cyclic three-edge line subgraphs,

$$Q = \{(e, f, g) | e, f, g \in G_{\mathbf{d}}, e \neq f, f \neq g, e \neq g, f = (u, v), e = (u, \bullet), g = (\bullet, v)\},$$

we find

$$\Delta^2_{p_r} = \sum_{(e,f,g)\in Q} t_e t_g (1 - t_f) = \sum_{(e,f,g)\in Q} t_e t_g - \sum_{e,f,g\in Q} t_e t_g t_f = Y_1 - Y_2.$$

Vu's inequality will be applied to Y_1 and Y_2 separately. To upper bound the expected value of Y_1, we need an upper bound on the size of Q. Given f, the source of e and the target of g are fixed. Hence there are at most $d_{\max}^2$ triples in Q with a fixed edge

f. As f may be any edge, $|Q| \leq md_{\max}^2$. Together with $\mathbb{E}[t_e t_g] = q_r^2$ this implies that $\mathbb{E}[Y_1] \leq md_{\max}^2 q_r^2$. We differentiate Y_1 with respect to $t_{\widetilde{e}}$, to obtain:

$$\sum_{\substack{(e,f,g)\in Q\\ e=\widetilde{e}}} t_g + \sum_{\substack{(e,f,g)\in Q\\ g=\widetilde{e}}} t_e.$$

Since

$$\sum_{\substack{(e,f,g)\in Q\\ e=\widetilde{e}}} 1 \leq d_{\max}^2 \quad \text{and} \quad \sum_{\substack{(e,f,g)\in Q\\ g=\widetilde{e}}} 1 \leq d_{\max}^2,$$

we have $\mathbb{E}[\partial_{t_{\widetilde{e}}} Y_1] \leq 2d_{\max}^2 q_r$, and since Y_1 is a polynomial of degree 2 with all coefficients equal to 1, all second derivatives are at most 1. Together, these observations yield:

$$\mathbb{E}_0[Y_1] \leq \max\left(1, 2d_{\max}^2 q_r, md_{\max}^2 q_r^2\right), \quad \mathbb{E}_1[Y_1] \leq \max\left(1, 2d_{\max}^2 q_r\right) \quad \text{and} \quad \mathbb{E}_2[Y_1] \leq 1.$$

Similar to (i) it can be shown that $\lambda = \lambda_i$ and

$$\mathcal{E}_0 = 9\lambda_i^2 + 2md_{\max}^2 q_r^2, \quad \mathcal{E}_1 = 9\lambda_i + 2d_{\max}^2 q_r \quad \text{and} \quad \mathcal{E}_2 = 1,$$

fulfil the conditions of Theorem 2.13. Applying Vu's inequality and assuming $c \geq 8 \cdot 9c_2$, we thus obtain

$$\mathbb{P}\left[|Y_1 - \mathbb{E}[Y_1]| \geq \frac{\beta_r(\lambda_i)}{8}\right] \leq e^{-\Omega(\lambda_i)}.$$

Moving on to Y_2, we see that $\mathbb{E}[Y_2] \leq md_{\max}^2 q_r^3$ as $|Q| \leq md_{\max}^2$ and $\mathbb{E}[t_e t_f t_g] = q_r^3$. Differentiating Y_2 to with respect $t_{\widetilde{e}}$, we obtain

$$\sum_{\substack{(e,f,g)\in Q\\ e=\widetilde{e}}} t_f t_g + \sum_{\substack{(e,f,g)\in Q\\ f=\widetilde{e}}} t_e t_g + \sum_{\substack{(e,f,g)\in Q\\ g=\widetilde{e}}} t_e t_f.$$

This implies that $\mathbb{E}[\partial_{t_{\widetilde{e}}} Y_1] \leq 3d_{\max}^2 q_r$. Differentiating Y_2 to with respect $t_{\widetilde{e}}$ and $t_{\widetilde{f}}$ for $\widetilde{e} \neq \widetilde{f}$, we obtain

$$\sum_{\substack{(e,f,g)\in Q\\ e=\widetilde{e}\\ f=\widetilde{f}}} t_g + \sum_{\substack{(e,f,g)\in Q\\ e=\widetilde{e}\\ g=\widetilde{f}}} t_f + \sum_{\substack{(e,f,g)\in Q\\ f=\widetilde{e}\\ g=\widetilde{f}}} t_e + \sum_{\substack{(e,f,g)\in Q\\ f=\widetilde{e}\\ e=\widetilde{f}}} t_g + \sum_{\substack{(e,f,g)\in Q\\ g=\widetilde{e}\\ e=\widetilde{f}}} t_f + \sum_{\substack{(e,f,g)\in Q\\ g=\widetilde{e}\\ f=\widetilde{f}}} t_e.$$

In each of the sums, there is freedom to choose only one edge. As the source, the target or both are fixed for this edge, each summation is upper bounded by $d_{\max} q_r$. According to the definition of Q, at most two of the summations are non-zero, implying that $\mathbb{E}\left[\partial_{t_{\widetilde{e}}} \partial_{t_{\widetilde{f}}} Y_2\right] \leq 2d_{\max} q_r$. As Y_2 is a polynomial of degree 3 and all of its coefficients are 1, any third order partial derivative of Y_2 can be at most 1. We thus find:

$$\mathbb{E}_0[Y_2] \leq \max\left(1, 2d_{\max} q_r, 3d_{\max}^2 q_r^2, md_{\max}^2 q_r^3\right), \quad \mathbb{E}_1[Y_2] \leq \max\left(1, 2d_{\max} q_r, 3d_{\max}^2 q_r^2\right),$$
$$\mathbb{E}[Y_2] \leq \max(1, 2d_{\max} q_r) \text{ and} \quad \mathbb{E}_3[Y_2] \leq 1.$$

Vu's inequality is applied to Y_2 using $\lambda = \lambda_i$ and

$$\mathcal{E}_0 = 85\lambda_i^3 + 3md_{\max}^2 q_r^3, \quad \mathcal{E}_1 = 85\lambda_i^2 + 3d_{\max}^2 q_r^2, \quad \mathcal{E}_2 = 17\lambda_i + 2d_{\max} q_r \quad \text{and} \quad \mathcal{E}_3 = 1,$$

to obtain

$$\mathbb{P}\left[|Y_2 - \mathbb{E}[Y_2]| \geq 85c_3\sqrt{\lambda_i\left(\lambda_i^2 + d_{\max}^2 q_r^2\right)\left(\lambda_i^3 + md_{\max}^2 q_r^3\right)}\right] \leq e^{-\Omega(\lambda_i)}.$$

If we choose c large enough, this implies that

$$\mathbb{P}\left[\left|\Delta_{p_r}^2 - \mathbb{E}\left[\Delta_{p_r}^2\right]\right| \geq \frac{\beta_r(\lambda_i) + \gamma_r(\lambda_i)}{8}\right] \leq e^{-\Omega(\lambda_i)}.$$

Next, remark that

$$\begin{aligned}\left|\Delta_{p_r}^2 - \mathbb{E}\left[\Delta_{p_r}^2\right]\right| &= |Y_1 - Y_2 - \mathbb{E}[Y_1] + \mathbb{E}[Y_2]| \leq |Y_1 - \mathbb{E}[Y_1]| + \mathbb{E}[Y_2] \\ &\leq |Y_1 - \mathbb{E}[Y_1]| + md_{\max}^2 q_r^3 = |Y_1 - \mathbb{E}[Y_1]| + \frac{\nu_r}{8}.\end{aligned}$$

This implies that there also holds

$$\mathbb{P}\left[\left|\Delta_{p_r}^2 - \mathbb{E}\left[\Delta_{p_r}^2\right]\right| \geq \frac{\beta_r(\lambda_i) + \nu_r}{8}\right] \leq e^{-\Omega(\lambda_i)},$$

completing the proof.

(iii) To prove that $\mathbb{P}\left[\left|\frac{\Lambda_{p_r}^{1\ -}\Lambda_{p_r}^{1\ +} - \Lambda_{p_r}^2}{4m} - \frac{\mathbb{E}\left[\Lambda_{p_r}^{1\ -}\Lambda_{p_r}^{1\ +} - \Lambda_{p_r}^2\right]}{4m}\right| \geq \frac{\beta_r(\lambda_i)}{8}\right] \leq e^{-\Omega(\lambda_i)}$, Vu's inequality is applied to $\frac{\Lambda_{p_r}^{1\ +}\Lambda_{p_r}^{1\ -}}{d_{\max}^2}$ and $\frac{\Lambda_{p_r}^2}{d_{\max}^2}$ separately. The construction is almost identical to the proofs of (i) and (ii). First consider

$$\begin{aligned}\frac{\Lambda_{p_r}^{1\ +}\Lambda_{p_r}^{1\ -}}{d_{\max}^2} &= \frac{\sum_{i=1}^n d_i^{-(r)} d_i^- \sum_{i=1}^n d_i^{+(r)} d_i^+}{d_{\max}^2} = \left(\sum_{e=(u,v)\in G_{\mathbf{d}}} \frac{d_u^-}{d_{\max}} t_e\right)\left(\sum_{f=(w,z)\in G_{\mathbf{d}}} \frac{d_z^+}{d_{\max}} t_f\right) \\ &= \left(\sum_{e=(u,v)\in G_{\mathbf{d}}} \frac{d_u^- d_v^+}{d_{\max}^2} t_e^2\right) + \sum_{\substack{e=(u,v)\in G_{\mathbf{d}} \\ f=(w,z)\in G_{\mathbf{d}} \\ e\neq f}} \frac{d_u^- d_z^+}{d_{\max}^2} t_e t_f = Z_1 + Z_2.\end{aligned}$$

Start with Z_1. This is a polynomial of degree one, as for a Bernoulli variable there holds $t_e^2 = t_e$. Since its coefficients are at most 1, it is clear that any first order partial derivative of Z_1 is upper bounded by 1. The expected value of Z_1 is upper bounded by mq_r. This implies that,

$$\mathbb{E}_0[Z_1] \leq \max(1, mq_r) \quad \text{and} \quad \mathbb{E}_1[Z_1] \leq 1.$$

Hence, $\mathcal{E}_0 = mq_r + \lambda_i$ and $\mathcal{E}_1 = 1$, satisfy the constraints of Theorem 2.13 with $\lambda = \lambda_i$. Applying this theorem we find

$$\mathbb{P}\left[|Z_1 - \mathbb{E}[Z_1]| \geq c_1\sqrt{\lambda_i(\lambda_i + mq_r)}\right] \leq e^{-\Omega(\lambda_i)}.$$

Next, consider Z_2. This is a sum over all pairs of distinct edges, hence it contains fewer than m^2 terms. Combining this with $\frac{d_u^- d_z^+}{d_{\max}^2} \leq 1$ and $\mathbb{E}\left[t_e t_f\right] = q_r^2$, we find that $\mathbb{E}\left[Z_2\right] \leq m^2 q_r^2$. Taking the partial derivative with respect to a variable t_g and writing $g = (i,j)$ leads to

$$\sum_{\substack{f=(w,z)\in G_{\mathbf{d}}\\ f\neq g}} \frac{d_i^- d_z^+}{d_{\max}^2} t_f + \sum_{\substack{e=(u,v)\in G_{\mathbf{d}}\\ e\neq g}} \frac{d_u^- d_j^+}{d_{\max}^2} t_e.$$

Each term of the summations is upper bounded by q_r. Each summation contains $m-1$ terms. Thus we find: $\mathbb{E}\left[\partial_{t_g} Z_2\right] \leq 2mq_r$. As Z_2 is a second order polynomial with coefficients upper bounded by 1, all second order partial derivatives will be at most 1. Combining these observations we find:

$$\mathbb{E}_0\left[Z_2\right] \leq \max\left(1, 2mq_r, m^2 q_r^2\right), \quad \mathbb{E}_1\left[Z_2\right] \leq \max\left(1, 2mq_r\right) \quad \text{and} \quad \mathbb{E}_2\left[Z_2\right] \leq 1.$$

Similar to the proof of (i) it can be shown that $\lambda = \lambda_i$ and

$$\mathcal{E}_0 = 9\lambda_i^2 + 2m^2 q_r^2, \quad \mathcal{E}_1 = 9\lambda_i + 2mq_r \quad \text{and} \quad \mathcal{E}_2 = 1,$$

satisfy the constraints of Vu's concentration inequality, which gives

$$\mathbb{P}\left[\left|Z_2 - \mathbb{E}\left[Z_2\right]\right| \geq 9c_2\sqrt{\lambda_i\left(\lambda_i + mq_r\right)\left(\lambda_i^2 + m^2 q_r^2\right)}\right] \leq e^{-\Omega(\lambda_i)}.$$

As $\sqrt{\lambda_i\left(\lambda_i + mq_r\right)} \leq \sqrt{\lambda_i\left(\lambda_i + mq_r\right)\left(\lambda_i^2 + m^2 q_r^2\right)}$ and $\frac{\Lambda_{p_r}^{1\ +}\Lambda_{p_r}^{1\ -}}{4m} = \frac{d_{\max}^2}{4m}\left(Z_1 + Z_2\right)$, we obtain

$$\mathbb{P}\left[\left|\frac{\Lambda_{p_r}^{1\ -}\Lambda_{p_r}^{1\ +}}{4m} - \frac{\mathbb{E}\left[\Lambda_{p_r}^{1\ -}\Lambda_{p_r}^{1\ +}\right]}{4m}\right| \geq \frac{d_{\max}^2}{m}(9c_2 + c_1)\sqrt{\lambda_i\left(\lambda_i + mq_r\right)\left(\lambda_i^2 + m^2 q_r^2\right)}\right] \leq e^{-\Omega(\lambda_i)}.$$

Pulling this factor $\frac{d_{\max}^2}{m}$ inside the root and taking $c > 8\left(c_1 + 9c_2\right)$, we also find

$$\mathbb{P}\left[\left|\frac{\Lambda_{p_r}^{1\ -}\Lambda_{p_r}^{1\ +}}{4m} - \frac{\mathbb{E}\left[\Lambda_{p_r}^{1\ -}\Lambda_{p_r}^{1\ +}\right]}{4m}\right| \geq \frac{c}{8}\sqrt{\lambda_i\left(\lambda_i + d_{\max}^2 q_r\right)\left(\lambda_i^2 + m d_{\max}^2 q_r^2\right)}\right] \leq e^{-\Omega(\lambda_i)}.$$

Next, we consider

$$\frac{\Lambda_{p_r}^2}{d_{\max}^2} = \sum_{i=1}^{n} \frac{d_i^{-(r)} d_i^- d_i^{+(r)} d_i^+}{d_{\max}^2} = \sum_{i=1}^{n} \frac{d_i^- d_i^+}{d_{\max}^2}\left(\sum_{e=(i,\bullet)\in G_{\mathbf{d}}} t_e\right)\left(\sum_{f=(\bullet,i)\in G_{\mathbf{d}}} t_f\right).$$

Note that this is the same expression as for $\Delta_{p_r}^1$ where the coefficient of each term is replaced by $\frac{\Lambda_{p_r}^2}{d_{\max}^2}$. Hence using the same argument as for (i) we obtain

$$\mathbb{P}\left[\left|\frac{\Lambda_{p_r}^2}{4m} - \frac{\mathbb{E}\left[\Lambda_{p_r}^2\right]}{4m}\right| \geq 9c_2\frac{d_{\max}^2}{4m}\sqrt{\lambda_i\left(\lambda_i + q_r d_{\max}\right)\left(\lambda_i^2 + m d_{\max} q_r^2\right)}\right] \leq e^{-\Omega(\lambda_i)}.$$

Again pulling $\frac{d_{\max}^2}{m}$ inside the square root, we find

$$\mathbb{P}\left[\left|\frac{\Lambda^1_{p_r}{}^- \Lambda^1_{p_r}{}^+ - \Lambda^2_{p_r}}{4m} - \frac{\mathbb{E}\left[\Lambda^1_{p_r}{}^- \Lambda^1_{p_r}{}^+ - \Lambda^2_{p_r}\right]}{4m}\right| \geq 9c_2\sqrt{\lambda_i\left(\lambda_i + d_{\max}^2 q_r\right)\left(\lambda_i^2 + m d_{\max}^2 q_r^2\right)}\right] \leq e^{-\Omega(\lambda_i)}.$$

Since $\beta = c\sqrt{\lambda_i\left(\lambda_i + d_{\max}^2 q_r\right)\left(\lambda_i^2 + m d_{\max}^2 q_r^2\right)}$, this completes the proof by taking $c > 8(18c_2 + c_1)$.

(iv) This argument is exactly the same as for (ii), as there holds

$$\frac{\Lambda^3_{p_r}}{d_{\max}^2} = \sum_{\substack{(e,f,g)\in Q \\ e=(u,v)}} \frac{d_u^+ d_v^-}{d_{\max}^2} t_e\left(1 - t_f\right) t_g.$$

Hence we obtain $\mathbb{P}\left[\left|\frac{\Lambda^3_{p_r}}{2m} - \frac{\mathbb{E}\left[\Lambda^3_{p_r}\right]}{2m}\right| \geq \frac{d_{\max}^2}{2m}\frac{\min(\beta_r(\lambda_i)+\gamma_r(\lambda_i),\beta_r(\lambda_i)+\nu_r)}{8}\right] \leq e^{-\Omega(\lambda_i)}$, and since $\frac{d_{\max}^2}{m} = o(1)$, this completes the proof.

□

Combining all inequalities from the statement of Lemma (2.14), we find that

$$\mathbb{P}\left[\left|\Psi_{p_r} - \mathbb{E}\left[\Psi_{p_r}\right]\right| \geq \frac{4\beta_r(\lambda_i) + 2\min(\nu_r, \gamma_r(\lambda_i))}{8}\right] \leq e^{-\Omega(\lambda_i)},$$

for all $i \in \{0, 1, \ldots, L\}$ and $0 \leq r \leq m-1$. By definition of ψ_r this shows equation (38) and hence it proves Lemma 2.12. This completes the proofs of Lemma 2.8 (a) and 2.9 (a).

Next, we prove Lemma 2.8 (b) and 2.9 (b). This requires the following Lemma.

Lemma 2.15. *For all $i \in \{1, 2, \ldots, L\}$ and $\mathcal{N} \in A_i \setminus A_{i-1}$ there holds*

$$\sum_{r=0}^{m-1} \frac{\max\left(\Psi_r\left(\mathcal{N}\right) - \psi_r, 0\right)}{(m-r)^2 - \Psi_r\left(\mathcal{N}\right)} = o\left(\lambda_i\right).$$

Furthermore for all $\mathcal{N} \in A_0$ there holds

$$\sum_{m-r=\lambda_0\omega}^{m} \frac{\max\left(\Psi_r\left(\mathcal{N}\right) - \psi_r, 0\right)}{(m-r)^2 - \Psi_r\left(\mathcal{N}\right)} = o\left(1\right).$$

Proof. The first claim follows by changing the summation $\sum_{m-r=2}^{2m-2}$ into $\sum_{m-r=1}^{m}$ in the proof of Lemma 15(b) [15]. The second claim follows by applying a similar change to the proof of Lemma 18 [15]. □

We will now determine an upper bound on $f(\mathcal{N})$ for all $\mathcal{N} \in S^*\left(\mathcal{M}\right)$. According to the definition of $S^*\left(\mathcal{M}\right)$, $\Psi_r\left(\mathcal{N}\right) \leq \left(1 - \frac{\tau}{4}\right)(m-r)^2$ holds for all $0 \leq r \leq m-1$. Therefore,

$$f(\mathcal{N}) = \prod_{r=0}^{m-1}\left(1 + \frac{\Psi_r\left(\mathcal{N}\right) - \psi_r}{(m-r)^2 - \Psi_r\left(\mathcal{N}\right)}\right) \leq \prod_{r=0}^{m-1}\left(1 + \frac{4\max\left(\Psi_r\left(\mathcal{N}\right) - \psi_r, 0\right)}{\tau(m-r)^2}\right).$$

Using $1+x \leq e^x$ the latter inequality becomes

$$f(\mathcal{N}) \leq e^{\sum_{r=0}^{m-1} \frac{4\max(\Psi_r(\mathcal{N})-\psi_r,0)}{\tau(m-r)^2}}. \tag{40}$$

Let us consider $\mathcal{N} \in A_i \setminus A_{i-1}$ for $i \in \{1, 2, \ldots, L\}$ and apply Lemma 2.15 to equation (40), to obtain:

$$f(\mathcal{N}) \leq e^{o(\lambda_i)}.$$

This completes the proof of Lemma 2.8 (b).

It remains to prove Lemma 2.9 (b). As $A_\infty \subset S^*(\mathcal{M})$ we have:

$$f(\mathcal{N}) \leq \prod_{r=0}^{m-d_{\max}^2} \left(1 + \frac{4\max(\Psi_r(\mathcal{N}) - \psi_r, 0)}{\tau(m-r)^2}\right) \prod_{r=m-d_{\max}^2+1}^{m-1} \frac{(m-r)^2 - \psi_r}{(m-r)^2 - \Psi_r(\mathcal{N})}.$$

Since $0 < \Psi_r(\mathcal{N})$ and $\psi_r < (m-r)^2$, we further have:

$$f(\mathcal{N}) \leq \left(d_{\max}^4\right)^{d_{\max}^2} \prod_{r=0}^{m-d_{\max}^2} \left(1 + \frac{4\Psi_r(\mathcal{N})}{\tau(m-r)^2}\right).$$

From Lemma 2.3 it follows that $\Psi_r = \Delta_r + \Lambda_r \leq 2(m-r)d_{\max}^2$, which, when inserted in the latter inequality, gives:

$$f(\mathcal{N}) \leq \left(d_{\max}^4\right)^{d_{\max}^2} \prod_{r=0}^{m-d_{max}^2} \left(1 + \frac{8d_{\max}^2}{\tau(m-r)}\right).$$

Using $(1+x) \leq e^x$, we find:

$$f(\mathcal{N}) \leq e^{4d_{\max}^2 \ln(d_{\max}) + \frac{8}{\tau}\sum_{i=d_{\max}^2}^{m} i^{-1} d_{\max}^2} \leq e^{4d_{\max}^2 \ln(d_{\max}) + \frac{8}{\tau}\ln(m) - \frac{8}{\tau}\ln(d_{\max}^2)},$$

and since $\tau \leq \frac{1}{3}$ and $m \leq n d_{\max}$, we have:

$$\begin{aligned} f(\mathcal{N}) &\leq e^{4d_{\max}^2 \ln(d_{\max}) + 24\ln(m)} \leq e^{4d_{\max}^2 \ln(d_{\max}) + 24\ln(n d_{\max})} \\ &\leq e^{24 d_{\max}^2 \ln\left(n d_{\max}^2\right)} \leq e^{24 d_{\max}^2 \ln\left(n^3\right)} = e^{72 d_{\max}^2 \ln(n)}. \end{aligned}$$

This proves Lemma 2.9 (b), completes the proofs of Lemma's 2.8 and 2.9, and therefore, completes the proof of the asymptotic estimate (20).

Proving equation (21) The next step is showing that equation (21) holds. To this end, we first prove the following Lemma.

Lemma 2.16. *For all* $1 \leq j \leq K$

(a) $\mathbb{P}[\mathcal{N} \in B_j \setminus B_{j-1}] \leq e^{-\Omega\left(2^{j/2}\ln(n)\right)}$*;*

(b) For all $\mathcal{N} \in B_j \setminus B_{j-1}$*,* $f(\mathcal{N}) \leq e^{\mathcal{O}\left(2^j\right)}$*.*

Proof. (a) The probability that $\mathcal{N} \in B_j \setminus B_{j-1}$ is upper bounded by the probability that $\mathcal{N} \in B_{j-1}^c := S(\mathcal{M}) \setminus B_{j-1}$. Hence if we show that

$$\mathbb{P}\left[\mathcal{N} \in B_j^c\right] \leq e^{-\Omega\left(2^{j/2}\ln(n)\right)},$$

the claim is proven. Remark that $B_{j-1}^c \subset \left\{\mathcal{N} \in S(\mathcal{M}) \,|\, \exists\, r, \text{ s.t. } m - r \leq \omega\lambda_0 \text{ and } \Psi_r \geq 2^{j-1}\right\}$. Therefore, we need to consider only those r for which $m - r \leq \omega\lambda_0$.

Note that

$$\mathbb{P}\left[\Psi_r \geq 2^{j-1} | r \geq m - \omega\lambda_0\right] \leq e^{-\Omega\left(2^{j/2}\ln(n)\right)} \tag{41}$$

is a stronger statement than the desired inequality. Indeed, using $\omega\lambda_0 \ll \ln(n)^2$ gives:

$$\begin{aligned}\mathbb{P}\left[\mathcal{N} \in B_{j-1}^c\right] &\leq \ln^2(n)\mathbb{P}\left[\Psi_r \geq 2^{j-1}\right] \leq \ln^2(n)e^{-\Omega\left(2^{j/2}\ln(n)\right)} \\ &= e^{-\Omega\left(2^{j/2}\ln(n)\right)+2\ln(\ln(n))} = e^{-\Omega\left(2^{j/2}\ln(n)\right)}.\end{aligned}$$

We will therefore prove inequality (41) instead. Fix an arbitrary r such that $m - r < \omega\lambda_0$ and assume that $\Psi_r \leq 2^{j-1}$. Then by definition of Ψ_r and applying Lemma 2.3, we have

$$\Delta_r \geq 2^{j-1} - \frac{d_{\max}^2 m}{2} q_r^2.$$

Since $m - r \leq \omega\lambda_0 < 2^{j-1}\omega\lambda_0$ and $d_{\max}^2\omega^2\lambda_0^2 < m$,

$$\begin{aligned}\Delta_r &\geq 2^{j-1} - \frac{2^{j-1}d_{\max}^2}{2m}\omega^2\lambda_0^2. \\ &\geq 2^{j-1} - \frac{2^{j-1}}{2} = 2^{j-2}.\end{aligned}$$

The remainder of the proof is similar to the proof of Lemma 2.11 wherein equation (37) is replaced by

$$2^{j-2} \leq \Delta_{p_r} \leq \sum_{u \in V} d_{G_q}^+(u) \sum_{v \in N_0(u)} d_{G_q}^-(v).$$

This inequality can be shown to imply one of the following statements holds true:

(a) G_q has more than $2^{j/2-1}$ edges;

(b) for some $u \in V$ there holds $\sum_{v \in N_0(u)} d_{G_q}^-(v) \geq 2^{j/2-1}$.

Indeed, the probability that either of those statements holds, is upper bounded by $e^{-\Omega\left(2^{j/2}\ln(n)\right)}$, by using the same argument as in the proof of Lemma 2.11. Since r is arbitrary this shows that $\mathbb{P}\left[\Psi_r \geq 2^{j-1}\right] \leq e^{-\Omega\left(2^{j/2}\ln(n)\right)}$ for all r such that $m - r < \omega\lambda_0$, completing the proof.

(b) Since $B_j \subset S^*(\mathcal{M})$ for all $1 \leq j \leq K$, inequality (40) gives

$$f(\mathcal{N}) \leq e^{\sum_{r=0}^{m-1} \frac{4\max(\Psi_r(\mathcal{N})-\psi_r,0)}{\tau(m-r)^2}},$$

for all $\mathcal{N} \in B_j \setminus B_{j-1}$. According the definition of B_j, we have

$$\sum_{m-r=1}^{\omega\lambda_0} \frac{\max\left(\Psi_r\left(\mathcal{N}\right) - \psi_r, 0\right)}{(m-r)^2} \leq \sum_{m-r=1}^{\omega\lambda_0} \frac{2^j}{(m-r)^2} = \mathcal{O}\left(2^j\right),$$

and since $B_j \subset A_0$ the second statement from Lemma 2.15 can be applied, giving:

$$\sum_{m-r=\omega\lambda_0}^{m} \frac{4\max\left(\Psi_r\left(\mathcal{N}\right) - \psi_r, 0\right)}{\tau(m-r)^2} = o(1).$$

Hence for all $\mathcal{N} \in B_j$ it holds $f(\mathcal{N}) \leq e^{\mathcal{O}\left(2^j\right)+o(1)} = e^{\mathcal{O}\left(2^j\right)}$.

□

Now, we give a proof of asymptotical estimate (21). Lemma 2.16 implies that for all $B_j \setminus B_{j-1}$

$$\mathbb{E}\left[f(\mathcal{N})\, \mathbb{1}_{B_j \setminus B_{j-1}}\right] \leq e^{-\Omega\left(2^{j/2}\ln(n)\right)} e^{\mathcal{O}\left(2^j\right)}.$$

Recall that $j \leq K$, and, in combination with equation (33), this yields $2^{\frac{j-1}{2}} \leq \ln(n)$. Hence there holds

$$\mathbb{E}\left[f(\mathcal{N})\, \mathbb{1}_{\mathcal{B}}\right] = \sum_{j=1}^{K} \mathbb{E}\left[f(\mathcal{N})\, \mathbb{1}_{B_j \setminus B_{j-1}}\right] \leq \sum_{j=1}^{K} e^{-\Omega\left(2^{j/2}\ln(n)\right)} e^{\mathcal{O}\left(2^j\right)} = o(1),$$

proving equation (21).

Proving equations (22) **and** (23) We bound the expected value of $f(\mathcal{N})$ for all $\mathcal{N} \in \mathcal{C}$. We, start with proving upper bound (22), for which it suffices to show that for all $\mathcal{N} \in \mathcal{C}$,

$$f(\mathcal{N}) \leq 1 + o(1).$$

As $\mathcal{C} \subset S^*\left(\mathcal{M}\right)$, in analogy to equation (40), there holds

$$\begin{aligned} f(\mathcal{N}) &= \prod_{r=0}^{m-1} \left(1 + \frac{\Psi_r\left(\mathcal{N}\right) - \psi_r}{(m-r)^2 - \Psi_r\left(\mathcal{N}\right)}\right) \\ &\leq \prod_{m-r=1}^{\lambda_0\omega} \left(1 + \frac{4\max\left(\Psi_r\left(\mathcal{N}\right) - \psi_r, 0\right)}{\tau(m-r)^2}\right) e^{\sum_{m-r=\lambda_0\omega+1}^{m} \frac{4\max\left(\Psi_r(\mathcal{N}) - \psi_r, 0\right)}{\tau(m-r)^2}}. \end{aligned}$$

Because $\mathcal{C} \subset A_0$, we obtain from Lemma 2.15 that

$$\sum_{m-r=\lambda_0\omega+1}^{m} \frac{4\max\left(\Psi_r\left(\mathcal{N}\right) - \psi_r, 0\right)}{\tau(m-r)^2} = o(1).$$

Also by definition of $\mathcal{C}$, $\Psi_r(\mathcal{N}) \leq 1$ for all $m-r \leq \omega\lambda_0$. Hence for all $\mathcal{N} \in \mathcal{C}$,

$$\begin{aligned} f(\mathcal{N}) &\leq \prod_{m-r=1}^{\lambda_0\omega} \left(1 + \frac{4}{\tau(m-r)^2}\right) e^{o(1)} \leq \left(1 + \mathcal{O}\left(\frac{4\lambda_0\omega}{\tau} \prod_{m-r=1}^{\lambda_0\omega} \frac{1}{(m-r)^2}\right)\right) e^{o(1)} \\ &\leq e^{o(1)}(1+o(1)) = 1+o(1), \end{aligned}$$

proving equation (22).

Next, we derive a lower bound on $\mathbb{E}\left[f(\mathcal{N})\,\mathbb{1}_{S^*(\mathcal{M})}\right]$. As $\mathcal{C} \subset S^*(\mathcal{M})$ this will prove equation (23). Take any ordering $\mathcal{N} \in S^*(\mathcal{M})$. Lemma 2.12 states that

$$\mathbb{P}\left[|\Psi_r(\mathcal{N}) - \psi_r| \geq 4\beta_r(\lambda_0) + 2\min(\gamma_r(\lambda_0), \nu_r)\right] \leq e^{-\Omega(\lambda_0)} < e^{-\ln(n)^{1+\delta}} = o(1), \tag{42}$$

holds for all r, such that $m - r \geq \omega\lambda_0$. Thus the probability that $|\Psi_r(\mathcal{N}) - \psi_r| \geq 4\beta_r(\lambda_0) + 2\min(\gamma_r(\lambda_0), \nu_r)$ holds for at least one r is small. Now consider an ordering $\mathcal{N} \in S^*(\mathcal{M})$ such that for all r with $m - r \geq \omega\lambda_0$ there holds

$$|\Psi_r(\mathcal{N}) - \psi_r| \leq 4\beta_r(\lambda_0) + 2\min(\gamma_r(\lambda_0), \nu_r). \tag{43}$$

Recall that $\mathcal{N} \in S^*(\mathcal{M})$ implies $\Psi_r(\mathcal{N}) \leq \left(1 - \frac{\tau}{4}\right)(m-r)^2$. Combining this with the definition of $f(\mathcal{N})$, we find:

$$\begin{aligned} f(\mathcal{N}) &\geq \prod_{m-r=\omega\lambda_0^3}^{m} \left(1 - \frac{\Psi_r(\mathcal{N}) - \psi_r}{(m-r)^2 - \Psi_r(\mathcal{N})}\right) \prod_{m-r=1}^{\omega\lambda_0^3+1} \left(1 - \frac{\psi_r}{(m-r)^2 - \Psi_r(\mathcal{N})}\right) \\ &\geq \prod_{m-r=\omega\lambda_0^3+1}^{m} \left(1 - \frac{4}{\tau} \frac{4\beta_r(\lambda_0) + 2\min(\gamma_r(\lambda_0), \nu_r)}{(m-r)^2}\right) \prod_{m-r=1}^{\omega\lambda_0^3} \left(1 - \frac{4}{\tau} \frac{\psi_r}{(m-r)^2}\right). \end{aligned}$$

From Lemma 2.15 and the definition of T_r, we find $\sum_{m-r=\omega\lambda_0^3+1}^{m} \frac{4}{\tau} \frac{4\beta_r(\lambda_0)+2\min(\gamma_r(\lambda_0),\nu_r)}{(m-r)^2} = o(1)$, which when combined with $1 - x \geq e^{-2x}$ for $0 \leq x \leq \frac{1}{2}$, gives

$$f(\mathcal{N}) \geq e^{-o(1)} \prod_{m-r=1}^{\omega\lambda_0^3} \left(1 - \frac{4}{\tau} \frac{\psi_r}{(m-r)^2}\right).$$

To approximate the remaining product, we apply Lemma 2.6 in combination with $1-x \geq e^{-2x}$ and asymptotical estimate $\lambda_0^3\omega d_{\max}^2 = o(m)$ to obtain:

$$f(\mathcal{N}) \geq e^{-2o(1)} \geq 1 - o(1).$$

Now, for each $\mathcal{N} \in S^*(\mathcal{M})$ we have shown that either $f(\mathcal{N}) \geq 1 - o(1)$ or that its probability is upper bounded by $o(1)$, which this completes the proof of equation (23). Remark that in fact we have proven

$$\mathbb{E}\left[f(\mathcal{N})\,\mathbb{1}_{S^*(\mathcal{M})}\right] \geq 1 - o(1).$$

Additionally, the proofs of equations (20)-(23) demonstrate the following corollary.

Corollary 2.17. *For a sufficiently large constant c, as used in the definition of λ_L, there holds:*

$$\mathbb{E}\left[\exp\left(\frac{1}{\tau^2} \sum_{r=0}^{m-1} \frac{\max(\Psi_r(\mathcal{N}) - \psi_r, 0)}{(m-r)^2}\right)\right] = 1 + o(1).$$

This corollary will be used to prove equation (24).

Proving equation (24)**.** This equation is the last bit that remains to prove equation (5). It concerns the expected value of $f(\mathcal{N})$ for the orderings in $S(\mathcal{M}) \setminus S^*(\mathcal{M})$. Equation (16) implies that for any $\mathcal{N} \in S(\mathcal{M}) \setminus S^*(\mathcal{M})$, there exists at least one $0 \le r \le m-1$ such that the inequality

$$\Psi_r(\mathcal{N}) \le \left(1 - \frac{\tau}{4}\right)(m-r)^2 \tag{44}$$

is violated. This inequality can only be violated for specific values of r. To determine these values, we assume that the above inequality is violated and investigate what are the implications for Δ_r. Recall that $\Psi_r = \Delta_r + \Lambda_r$. By using Lemma 2.3 to bound Λ_r, we obtain:

$$\Delta_r > \Psi_r - \frac{d_{\max}^2}{2m}(m-r)^2.$$

Since $d_{\max}^4 = o(m)$, there is such n_0 that for all $n > n_0$ there holds $\frac{d_{\max}^2}{m} < \frac{\tau}{2}$. Let $n > n_0$, then

$$\Delta_r > \Psi_r - \frac{\tau}{4}(m-r)^2.$$

Assuming the opposite inequality to (44) holds, this becomes:

$$\Delta_r > \left(1 - \frac{\tau}{2}\right)(m-r)^2. \tag{45}$$

Lemma 2.3 states that $\Delta_r \le (m-r)d_{\max}^2$, and hence, we deduce that $(m-r)\left(1-\frac{\tau}{2}\right) \le d_{\max}^2$, which is equivalent to

$$m - r \le \frac{2d_{\max}^2}{2-\tau}.$$

Therefore, inequality (44) can only be violated if $m - r \le \frac{2d_{\max}^2}{2-\tau}$. This allows us to partition

$$S(\mathcal{M}) \setminus S^*(\mathcal{M}) = \bigcup_{t=1}^{\frac{2d_{\max}^2}{2-\tau}} S_t(\mathcal{M}),$$

with $S_t(\mathcal{M})$ being the set of all orderings $\mathcal{N}$ violating inequality (44) with $r = m - t$ and not violating it for all $r < m - t$. To prove equation (24), it suffices to show that

$$\mathbb{E}\left[f(\mathcal{M})\,\mathbb{1}_{S_t}\right] \le \mathcal{O}\left(\frac{1}{m^{t\tau}}\right), \tag{46}$$

for all $t \in \{1, 2, \ldots, \frac{2d_{\max}^2}{2-\tau}\}$ as $\sum_{t=1}^{\infty} \frac{1}{m^{t\tau}} = o(1)$. We will now prove equation (46).

According to the definition of Ψ_r, we have $(m-r)^2 - \Psi_r = \sum_{(u,v)\in E_r} d_u^{+(r)} d_v^{-(r)} \left(1 - \frac{d_u^+ d_v^-}{2m}\right)$. For the algorithm to finish successfully, there must be at least $m-r$ suitable pairs left at each step r, implying that $(m-r)^2 - \Psi_r \ge (m-r)\left(1 - \frac{d_{\max}^2}{2m}\right)$. Therefore

$$\frac{(m-r)^2}{(m-r)^2 - \Psi_r} \le \frac{(m-r)}{1 - \frac{d_{\max}^2}{2m}} = (m-r)\left(1 + \mathcal{O}\left(\frac{d_{\max}^2}{2m}\right)\right),$$

and since $\frac{d_{\max}^4}{m} = o(1)$, we have: $\frac{(m-r)^2}{(m-r)^2-\Psi_r} \leq m - r + 1$ for $m - r \leq \frac{2d_{\max}^2}{2-\tau}$. Now we have that

$$\prod_{r=m-t}^{m-1} \frac{(m-r)^2 - \psi_r}{(m-r)^2 - \Psi_r} \leq \prod_{r=m-t}^{m-1} \frac{(m-r)^2}{(m-r)^2 - \Psi_r} \leq \prod_{r=m-t}^{m-1} m - r + 1 = (t+1)! \leq t^t(t+1).$$

In analogy to equation (40) it can also be shown that

$$\prod_{r=0}^{m-t} \frac{(m-r)^2 - \psi_r}{(m-r)^2 - \Psi_r} \leq \exp\left[\frac{4}{\tau} \sum_{r=0}^{m-1} \frac{\max(\Psi_r - \psi_r, 0)}{(m-r)^2}\right].$$

Combing these observations with inequality (44), which holds for all $r < m - t$, we find:

$$f(\mathcal{N})\, \mathbb{1}_{S_t} = \mathbb{1}_{S_t} \prod_{r=0}^{m-r} \frac{(m-r)^2 - \psi_r}{(m-r)^2 - \Psi_r} \leq \mathbb{1}_{S_t} \exp\left[\frac{4}{\tau} \sum_{r=0}^{m-1} \frac{\max(\Psi_r - \psi_r, 0)}{(m-r)^2}\right] t^t(t+1).$$

Next, we take the expected value of the above equation and apply Hölder's inequality to obtain:

$$\mathbb{E}\left[f(\mathcal{N})\, \mathbb{1}_{S_t}\right] \leq \mathbb{E}\left[\mathbb{1}_{S_t}\right]^{1-\tau} \mathbb{E}\left[\mathbb{1}_{S_t} \exp\left[\frac{4}{\tau^2} \sum_{r=0}^{m-1} \frac{\max(\Psi_r - \psi_r, 0)}{(m-r)^2}\right]\right]^{\tau} t^t(t+1).$$

Using Corollary 2.17 this becomes

$$\mathbb{E}\left[f(\mathcal{N})\, \mathbb{1}_{S_t}\right] \leq \mathbb{E}\left[\mathbb{1}_{S_t}\right]^{1-\tau} [1 + o(1)]\, t^t(t+1).$$

Hence, to prove equation (46), it remains to show that

$$\mathbb{P}\left[\mathcal{N} \in S_t\right]^{1-\tau} t^t(t+1) \leq [1 + o(1)]\, \frac{1}{m^{\tau t}}. \tag{47}$$

This requires an upper bound on $\mathbb{P}[\mathcal{N} \in S_t]$, which we derive in the following manner: As the first step, we show that if $\mathcal{N} \in S_t$, then $G_{\mathcal{N}_r}$ always contains a vertex with some special property. We use use the probability that such a vertex exists as an upper bound for $\mathbb{P}[\mathcal{N} \in S_t]$. Let us assume that $\mathcal{N} \in S_t$, fix $r = m - t$ and define $\Gamma(u) := \{v \in V | (u, v) \in G_{\mathcal{N}_r}\}$. By definition of Δ_r, this allows us to write

$$\Delta_r = \sum_{u \in V} d_u^{+(r)} \sum_{v \in \Gamma(u) \cup \{u\}} d_v^{-(r)} \quad \text{and} \quad (m-r)^2 = \sum_{u \in V} d_u^{+(r)} \sum_{v \in V} d_v^{-(r)}.$$

Because $\mathcal{N} \in S_t$, inequality (45) must hold. Inserting the above expressions for Δ_r and $(m-r)$ into this inequality yields:

$$\sum_{u \in V} d_u^{+(r)} \sum_{v \in \Gamma(u) \cup \{u\}} d_v^{-(r)} > \left(1 - \frac{\tau}{2}\right) \sum_{u \in V} d_u^{+(r)} \sum_{v \in V} d_v^{-(r)} > (1 - \tau) \sum_{u \in V} d_u^{+(r)} \sum_{v \in V} d_v^{-(r)},$$

which implies that there exists a vertex $u \in V$ such that

$$d_u^{+(r)} > 0 \quad \text{and} \quad \sum_{v \in \Gamma(u) \cup \{u\}} d_v^{-(r)} > (1 - \tau) \sum_{v \in V} d_v^{-(r)} = (1 - \tau)t. \tag{48}$$

Thus we have shown that if $\mathcal{N} \in S_t$, there must exists a vertex u obeying (48), and therefore, probability that $G_{\mathcal{N}_r}$ contains such a vertex u provides an upper bound for $\mathbb{P}[\mathcal{N} \in S_t]$. As the second step, we derive an upper bound on the probability that u obeys (48). Recall that $G_{\mathcal{N}_r}$ contains the first r edges of the ordering $\mathcal{N}$. Adding the remaining t edges of $\mathcal{N}$ completes $G_{\mathbf{d}}$. In this complement edge set, let l out of t edges have their target in $\Gamma(u) \cup \{u\}$, then $l = \sum_{v \in \Gamma(u) \cup \{u\}} d_v^{-(r)}$. Let $k := d_u^+ - |\Gamma(u)| = d_u^{+(r)}$. Inequality (48) holds if and only if $k \geq 1$ and $l \geq (1-\tau)t$. We derive an upper bound on the probability that $k \geq 1$ and $l \geq (1-\tau)t$ for a random ordering $\mathcal{N} \in S(\mathcal{M})$. That is to say we fix all m edges in the graph, but the order in which they are drawn $\mathcal{N}$ is a uniform random variable. To obtain a fixed value of k, exactly k of the d_u^+ edges with u as source must be in $\mathcal{N} \setminus \mathcal{N}_r$. Choosing these edges determines $\Gamma(u)$. To obtain the desired value of l, exactly l edges with target in $\Gamma(u) \cup \{u\}$ must be in $\mathcal{N} \setminus \mathcal{N}_r$. There are $\sum_{v \in \Gamma(u) \cup \{u\}} (d_v^- - 1) + d_u^-$ edges to choose from, since for each $v \in \Gamma(u)$ the edge with v as the target and u as the source is already in $\mathcal{N}_r$. The remaining $t - l - k$ edges that are not in $\mathcal{N}_r$ may be chosen freely amongst all the edges that do not have u as a source or an element of $\Gamma(u) \cup \{u\}$ as target. Thus the probability to get a specific combination of k and l is

$$\frac{\binom{d_u^+}{k}\binom{\sum_{v\in\Gamma(u)}(d_v^- -1)+d_u^-}{l}\binom{m-d_u^+-\sum_{v\in\Gamma(u)\cup}(d_v^- -1)-d_u^-}{t-l-k}}{\binom{m}{t}}.$$

We therefore write the upper bound for the probability that a randomly chosen vertex u satisfies (48) as

$$\sum_{k\geq 1, l\geq(1-\tau)t} \frac{\binom{d_u^+}{k}\binom{(d_u^+-k+1)d_{\max}}{l}\binom{m-d_u^+-\sum_{v\in\Gamma(u)}(d_v^- -1)-d_u^-}{t-l-k}}{\binom{m}{t}}.$$

For $\mathcal{N} \in S_t$ at least one vertex satisfies inequality (48), thus we have:

$$\mathbb{P}[\mathcal{N} \in S_t] \leq \sum_{u\in V} \sum_{k\geq 1, l\geq(1-\tau)t} \frac{\binom{d_u^+}{k}\binom{(d_u^+-k+1)d_{\max}}{l}\binom{m-d_u^+-\sum_{v\in\Gamma(u)}(d_v^- -1)-d_u^-}{t-l-k}}{\binom{m}{t}}.$$

Remark that $\binom{m}{k} \leq \frac{m^k}{k!}$, and since $t = \mathcal{O}\left(d_{\max}^2\right)$ and $\mathcal{O}\left(d_{\max}^4\right) = o(m)$ there holds:

$$\binom{m}{t} = [1 + o(1)]\, \frac{m^t}{t!}.$$

This gives

$$\begin{aligned}\mathbb{P}[S_t] &\leq \sum_{u\in V} \sum_{k\geq 1, l\geq(1-\tau)t} [1+o(1)]\, \frac{d_u^{+k}\left((d_u^+ - k + 1)(d_{\max})\right)^l m^{t-l-k} t!}{m^t k! l! (m-l-k)!} \\ &= \sum_{u\in V} \sum_{k\geq 1, l\geq(1-\tau)t} [1+o(1)]\, \frac{\left(\frac{d_u^+}{m}\right)^k \left(\frac{(d_u^+-k+1)d_{\max}}{m}\right)^l t!}{k! l! (m-l-k)!}.\end{aligned}$$

Finally, we approximate the sum over k and l. Since adding t edges completes the ordering, $\sum_{u\in V} d_u^{+(r)} = \sum_{u\in V} d_u^{-(r)} = t$. This implies that $k \in \{1, 2, \ldots t\}$ and that l is an integer in the interval $[(1-\tau)t, t]$. Thus this sum consists of at most $t\tau$ terms. Remark that, as $l, k \leq t = \mathcal{O}\left(d_{\max}^2\right) = \mathcal{O}\left(m^{1/2}\right)$, $\left(\frac{d_u^+}{m}\right) = \mathcal{O}\left(\frac{1}{m^{3/4}}\right)$ and $((d_u^+ - k + 1)(d_{\max})) = \mathcal{O}\left(\frac{1}{m^{1/2}}\right)$, the term inside the summation is maximal for $k = 1$ and $l = (1-\tau)\,t$. This gives:

$$\begin{aligned}
\mathbb{P}\left[S_t\right] &\leq [1+o(1)]\,\tau t \sum_{u\in V}\left(\frac{d_u^+}{m}\right)\left(\frac{d_u^+ d_{\max}}{m}\right)^{(1-\tau)t}\binom{t}{t\tau} \\
&\leq [1+o(1)]\,2^t t\left(\frac{d_{\max}^2}{m}\right)^{(1-\tau)t}\sum_{v\in V}\left(\frac{d_u^+}{m}\right) \\
&\leq [1+o(1)]\,2^t t\left(\frac{d_{\max}^2}{m}\right)^{(1-\tau)t}.
\end{aligned}$$

Here we used that $\tau \leq \frac{1}{3}$, $\binom{m}{k} \leq 2^m$ and $\sum_{u\in V} d_u^+ = m$. Plugging this into (47) yields:

$$\mathbb{P}\left[\mathcal{N}\in S_t\right]^{1-\tau} t^t(t+1) \leq [1+o(1)]\,t^t(t+1)\left(2^t t\left(\frac{d_{\max}^2}{m}\right)^{(1-\tau)t}\right)^{1-\tau}.$$

Since $t \leq \frac{2d_{\max}^2}{2-\tau}$, we have:

$$\mathbb{P}\left[\mathcal{N}\in S_t\right]^{1-\tau} t^t(t+1) \leq [1+o(1)]\,(t+1)t^{1-\tau}\left(\frac{2\cdot 2^{1-\tau}}{2-\tau}\frac{d_{\max}^{4-4\tau+2\tau^2}}{m^{1-2\tau+\tau^2}}\right)^t,$$

and since $\tau \leq \frac{1}{3}$, for any $x \geq 1$, $x^{1-\tau} \leq x$, we find:

$$\mathbb{P}\left[\mathcal{N}\in S_t\right]^{1-\tau} t^t(t+1) \leq [1+o(1)]\,(t+1)t\left(\frac{4}{2-\tau}\frac{d_{\max}^{4-4\tau+2\tau^2}}{m^{1-2\tau+\tau^2}}\right)^t.$$

Inserting the estimate $d_{\max} = \mathcal{O}\left(m^{1/4-\tau}\right)$ yields,

$$\mathbb{P}\left[\mathcal{N}\in S_t\right]^{1-\tau} t^t(t+1) \leq [1+o(1)]\,(t+1)t\left(\frac{4}{2-\tau}m^{-3\tau+3.5\tau^2-3\tau^3}\right)^t,$$

and using $t = o\left(m^{1/2}\right)$ and that $\frac{4}{2-\tau}$ is constant when m goes to infinity with n, we find:

$$\mathbb{P}\left[\mathcal{N}\in S_t\right]^{1-\tau} t^t(t+1) \leq [1+o(1)]\,o\left(m^{1/2}\right)\mathcal{O}\left(m^{-3\tau+3.5\tau^2-3\tau^3}\right)^t = \mathcal{O}\left(m^{-\tau t}\right).$$

This completes the proof of inequality (46) and hence it shows that equation (24) holds. This completes the prove of equation (15) and hence proves equation (4). Together with the results from Sections 2.1, and 2.3 this completes the proof of Theorem 2.1.

3 The probability of failure of Algorithm 1

Here we show that the probability the algorithm fails is $o(1)$. The proof is inspired by [15, Section 5]. If at step s, every pair of an unmatched in-stub with an unmatched out-stub is unsuitable – the algorithm fails. In this case, the algorithm will necessary create a self-loop or double edge when the corresponding edge is added to $G_{\mathcal{N}_s}$. First, we investigate at which steps $s \in \{0, 1, \ldots, m-1\}$ the algorithm can fail. Then, we derive an upper bound for the number of vertices that are left with unmatched stubs when the algorithm fails. For a given number of unmatched stubs, this allows us to determine the probability that the algorithm fails. Combining these results, we show that this probability is $o(1)$. The following lemma states that the algorithm has to be close to the end to be able to fail.

Lemma 3.1. *If Algorithm 1 fails at step s, then $m - s \leq d_{\max}^2$.*

Proof. At step s, there are $(m-s)^2$ pairs of unmatched stubs. If the algorithm fails at step s, all these pairs are unsuitable. The number of unsuitable pairs at step s is Δ_s. According to Lemma 2.3, $\Delta_s \leq d_{\max}^2(m-s)$. Therefore, if the algorithm fails at step s, there must hold $(m-s)^2 \leq d_{\max}^2(m-s)$. □

The number of vertices that have unmatched stubs when the algorithm fails is also bounded. Suppose a vertex $v \in V$ has unmatched in-stub(s) left when the algorithm fails. Since the number of unmatched in-stubs equals the number of unmatched out-stubs, this implies that there are also unmatched out-stubs. Because the algorithm fails, every pair of an unmatched in-stub and an unmatched out-stub induces either a double edge or self-loop. Hence, only v and vertices that are the source of an edge with v as a target can have unmatched out-stub(s). As v has at least one unmatched in-stub, there are at most $d_{\max} - 1$ edges with v as a target. Thus at most $d_{\max}$ vertices have unmatched out-stub(s). Symmetry implies that at most $d_{\max}$ vertices have unmatched in-stub(s) when a failure occurs.

Let $A_{d_{i_1}^{-\,(s)},\ldots,d_{i_{k^-}}^{-\,(s)},d_{j_1}^{+\,(s)},\ldots d_{j_{k^+}}^{+\,(s)}}$ be the event that the algorithm fails at step s with $v_{i_1}, \ldots, v_{i_{k^-}} \in V$ being the only vertices with unmatched in-stubs and $v_{j_1}, \ldots, v_{j_{k^+}}$ the only vertices having unmatched out-stubs. The amount of unmatched in-stubs (respectively out-stubs) of such a vertex i_l (j_l) is denoted by $d_{i_l}^{-(s)}$ ($d_{j_l}^{-(s)}$). Since k^- (respectively k^+) denotes the number of vertices with unmatched in-stubs (out-stubs) that are left, there holds $k^-, k^+ \leq d_{\max}$. This allows to write the probability that Algorithm 1 fails as

$$\mathbb{P}\left[\,\text{failure}\,\right] = \sum_{m-s=1}^{d_{\max}^2} \sum_{k^-,k^+=1}^{\max(m-s,d_{\max})} \sum_{i_1,\ldots,i_{k^-}=1}^{n} \sum_{j_1,\ldots,j_{k^+}=1}^{n} \mathbb{P}\left[A_{d_{i_1}^{-\,(s)},\ldots,d_{i_{k^-}}^{-\,(s)},d_{j_1}^{+\,(s)},\ldots d_{j_{k^+}}^{+\,(s)}}\right]. \tag{49}$$

The sum $\sum_{i_1,\ldots,i_{k^-}=1}^{n}$ is the sum over all possible subsets $B \subset \{1, 2, \ldots, n\}$ of size k^-, such that $\sum_{i \in B} d_i^{-(s)} = m - s$ and $\sum_{i \notin B} d_i^{-(s)} = 0$. The goal is to show that $\mathbb{P}\left[\,\text{failure}\,\right] = o(1)$, which we achieve by first determining an upper bound for $\mathbb{P}\left[A_{d_{i_1}^{-\,(s)},\ldots,d_{i_{k^-}}^{-\,(s)},d_{j_1}^{+\,(s)},\ldots d_{j_{k^+}}^{+\,(s)}}\right]$.

Lemma 3.2. *The probability of the event $A_{d_{i_1}^{-\,(s)},\ldots,d_{i_{k^-}}^{-\,(s)},d_{j_1}^{+\,(s)},\ldots d_{j_{k^+}}^{+\,(s)}}$ is upper bounded by*

$$e^{o(1)} d_{\max}^{2k^+k^- - 2k^{\pm}} \frac{\prod_{i \in K^+} {d_i^+}^{d_i^{+\,(s)}} \prod_{i \in K^-} {d_i^-}^{d_i^{-\,(s)}}}{m^{k^+k^- - k^{\pm}} m^{2(m-s)}} \binom{m-s}{d_{i_1}^{-\,(s)}, \ldots, d_{i_{k^-}}^{-\,(s)}} \binom{m-s}{d_{j_1}^{+\,(s)}, \ldots, d_{j_{k^+}}^{+\,(s)}}. \tag{50}$$

Proof. Let us define $K^{-} := \{i_1, i_2, \dots i_{k^-}\}$, $K^{+} := \{j_1, j_2, \dots j_{k^+}\}$, and $K^{\pm} := K^{-} \cap K^{+}$. When event $A_{d_{i_1}^{-(s)},\dots,d_{i_{k^-}}^{-(s)},d_{j_1}^{+(s)},\dots d_{j_{k^+}}^{+(s)}}$ occurs, the algorithm has constructed a graph $G_{\mathcal{M}_s}$ having the degree sequence $\widetilde{\mathbf{d}}$ with elements:

$$\widetilde{d_i^-} = \begin{cases} d_i^- & \text{if } i \notin K^- \\ d_i^- - d_i^{-(s)} & \text{if } i \in K^- \end{cases}, \quad \widetilde{d_i^-} = \begin{cases} d_i^+ & \text{if } i \notin K^+ \\ d_i^+ - d_i^{+(s)} & \text{if } i \in K^+ \end{cases}.$$

The probability of $A_{d_{i_1}^{-(s)},\dots,d_{i_{k^-}}^{-(s)},d_{j_1}^{+(s)},\dots d_{j_{k^+}}^{+(s)}}$ equals the number of graphs $G_{\mathcal{M}_s}$ that obey $\widetilde{\mathbf{d}}$ and lead to a failure multiplied by the probability that the algorithm constructs this partial graph. To construct an upper bound on the number of graphs obeying $\widetilde{\mathbf{d}}$ and leading to a failure, note that such a graph must contain the edge (i,j) for all $i \in K^+, j \in K^-, i \neq j$, and therefore, it must contain a subgraph obeying degree sequence $\overline{d_{K^-,K^+}}^{(s)}$, which is defined by:

$$\overline{d_i^-}^{(s)} := \begin{cases} d_i^- & \text{if } i \notin K^- \\ d_i^- - d_1^{-(s)} - k^+ & \text{if } i \in K^-, i \notin K^+ \\ d_i^- - d_1^{-(s)} - k^+ + 1 & \text{if } i \in K^-, i \in K^+ \end{cases}$$

and

$$\overline{d_i^+}^{(s)} := \begin{cases} d_i^+ & \text{if } i \notin K^+ \\ d_i^+ - d_1^{+(s)} - k^- & \text{if } i \in K^+, i \notin K^- \\ d_i^+ - d_1^{+(s)} - k^- + 1 & \text{if } i \in K^+, i \in K^- \end{cases}.$$

The number of graphs obeying the degree sequence $\overline{d_{K^-,K^+}}^{(s)}$ gives an upper bound for the number of partial graphs inducing event $A_{d_{i_1}^{-(s)},\dots,d_{i_{k^-}}^{-(s)},d_{j_1}^{+(s)},\dots d_{j_{k^+}}^{+(s)}}$. Denote by $\mathcal{L}(\mathbf{d})$ the space of simple graphs obeying the degree sequence $\mathbf{d}$. Theorem 2.1 implies that for any degree sequence d with $d_{\max} = \mathcal{O}\left(m^{1/4-\tau}\right)$ there holds

$$|\mathcal{L}(d)| \leq \frac{\prod_{r=0}^{m-1}(m-r)^2}{m! \prod_{i=1}^{n} d_i^+! \prod_{i=1}^{n} d_i^-!} e^{-\frac{\sum_{i=1}^n d_i^- d_i^+}{m} + \frac{\sum_{i=1}^n (d_i^-)^2 + (d_i^+)^2}{2m} - \frac{\sum_{i=1}^n (d_i^-)^2 \sum_{i=1}^n (d_i^+)^2}{4m^2} - \frac{1}{2} + o(1)}. \quad (51)$$

We apply this bound to the degree sequence $\overline{d_{K^-,K^+}}^{(s)}$. A graph obeying this degree sequence has $s - k^+k^- + k^{\pm}$ edges, with $k^{\pm} = |K^{\pm}|$. Thus we must show that $d_{\max} = \mathcal{O}\left((s - k^-k^+ + k^{\pm})^{1/4-\tau}\right)$. Combining the statement of Lemma 3.1 with $d_{\max}^4 = o(m)$ gives $s > 3d_{max}^2$ for $d_{\max} > 1$. Since $k^+k^- \leq d_{\max}^2$, we now find $m < 2(s - k^-k^+ + k^{\pm})$, that is $m = \mathcal{O}(s - k^-k^+ + k^{\pm})$, which implies that $d_{\max} = \mathcal{O}\left(m^{1/4-\tau}\right) = \mathcal{O}\left((s - k^-k^+k^{\pm})^{1/4-\tau}\right)$.

Thus we may apply inequality (51) to $\overline{d_{K^-,K^+}}^{(s)}$ to obtain

$$\left|\mathcal{L}\left(\overline{d_{K^-,K^+}}^{(s)}\right)\right| \leq \frac{(s - k^+k^- + k^{\pm})!}{\prod_{i=1}^n \overline{d_i^+}^{(s)}! \prod_{i=1}^n \overline{d_i^-}^{(s)}!}$$
$$\times \exp\left(\frac{\sum_{i=1}^n \left[(\overline{d_i^-}^{(s)})^2 + (\overline{d_i^+}^{(s)})^2 \right]}{2(s - k^+k^- + k^{\pm})} - \frac{\sum_{i=1}^n \overline{d_i^-}^{(s)} \overline{d_i^+}^{(s)}}{s - k^+k^- + k^{\pm}} - \frac{\sum_{i=1}^n (\overline{d_i^-}^{(s)})^2 \sum_{i=1}^n (\overline{d_i^+}^{(s)})^2}{4(s - k^+k^- + k^{\pm})^2} - \frac{1}{2} + o(1) \right).$$

Following the derivation in Sections 2.1 and 2.3 we find

$$\mathbb{P}_A\left(G_{\mathcal{M}_s}\right) = \frac{\prod_{i=1}^n d_i^+! \prod_{i=1}^n d_i^-!}{\prod_{i\in K^+} d_i^{+(s)}! \prod_{i\in K^-} d_i^{-(s)}!} \sum_{\mathcal{N}_s\in S(\mathcal{M}_s)} \mathbb{P}_A\left(\mathcal{N}_s\right)$$

$$= \frac{\prod_{i=1}^n d_i^+! \prod_{i=1}^n d_i^-!}{\prod_{i\in K^+} d_i^{+(s)}! \prod_{i\in K^-} d_i^{-(s)}!} s! \prod_{r=0}^{s-1} \frac{1}{(m-r)^2}$$

$$\times \exp\left(\frac{s\sum_{i=1}^n d_i^- d_i^+}{m^2} - \frac{s^2\sum_{i=1}^n\left[(d_i^-)^2+(d_i^+)^2\right]}{2m^3} + \frac{s\sum_{i=1}^n (d_i^-)^2\sum_{i=1}^n (d_i^+)^2}{4m^3} + \frac{s^2}{2m^2} + o(1)\right).$$

In the latter expression, the factor with factorials accounts for the number of different configurations leading to the same graph $G_{\mathcal{M}_s}$, which equals the number of permutations of the stub labels. However for $i \in K^-$ there are only $\frac{d_i^-!}{d_i^{-(s)}!}$ permutations of the labels of the in-stubs of v_i that lead to a different configuration. Remark that changing the label of an in-stub that remains unmatched with another in-stub that remains unmatched does not change the configuration. By the same argument for $i \in K^+$ there are only $\frac{d_i^+!}{d_i^{+(s)}!}$ ways to permute the labels of the out-stubs of v_i.

We can now determine

$$\mathbb{P}\left[A_{d_{i_1}^{-(s)},\ldots,d_{i_{k^-}}^{-(s)},d_{j_1}^{+(s)},\ldots d_{j_{k^+}}^{+(s)}}\right] \le \mathbb{P}\left[G_{\mathcal{M}_s}\right]\left|\mathcal{L}\left(\bar{d}_{k^-,k^+}^{(s)}\right)\right|.$$

First, we look at the product of the exponentials in the asymptotical approximations of $\mathbb{P}\left[G_{\mathcal{M}_s}\right]$ and $\left|\mathcal{L}\left(\bar{d}_{k^-,k^+}^{(s)}\right)\right|$, which after some transformations, and using that $m > s \ge m - d_{\max}^2$, becomes:

$$\exp\left(\frac{\sum_{i=1}^n\left[(\overline{d_i^-}^{(s)})^2+(\overline{d_i^+}^{(s)})^2\right]}{2\left(s-k^+k^-+k^\pm\right)} - \frac{\sum_{i=1}^n \overline{d_i^-}^{(s)}\overline{d_i^+}^{(s)}}{s-k^+k^-+k^\pm} - \frac{\sum_{i=1}^n(\overline{d_i^-}^{(s)})^2\sum_{i=1}^n(\overline{d_i^+}^{(s)})^2}{4\left(s-k^+k^-+k^\pm\right)^2} - \frac{1}{2} + o(1)\right)$$

$$= \exp\left(\frac{s}{m}\mathcal{O}\left(d_{\max}\right) + \frac{s}{m}\mathcal{O}\left(d_{\max}^2\right) + o(1)\right)\exp\left(-\mathcal{O}\left(d_{\max}\right) - \mathcal{O}\left(d_{\max}^2\right) + o(1)\right) = e^{o(1)}.$$

By using the latter estimate, we obtain

$$\mathbb{P}\left[A_{d_{i_1}^{-(s)},\ldots,d_{i_{k^-}}^{-(s)},d_{j_1}^{+(s)},\ldots d_{j_{k^+}}^{+(s)}}\right] \le \mathbb{P}\left[G_{\mathcal{M}_s}\right]\left|\mathcal{L}\left(\overline{d_{K^-,K^+}}^{(s)}\right)\right|$$

$$\le e^{o(1)}\frac{\prod_{i\in K^+} d_i^+! \prod_{i\in K^-} d_i^-! \prod_{i\in K^+, i\in K^-}\left(d_i^+ - d_i^{+(s)} - k^-\right)\left(d_i^- - d_i^{-(s)} - k^+\right)}{\prod_{i\in K^+}\left(d_i^+ - d_i^{+(s)} - k^-\right)! d_i^{+(s)}! \prod_{i\in K^-}\left(d_i^- - d_i^{-(s)} - k^+\right)! d_i^{-(s)}!}$$

$$\times \frac{(s-k^+k^-+k^\pm)!s!(m-s)!(m-s)!}{m!m!}$$

$$\le e^{o(1)} d_{\max}^{2k^+k^- - 2k^\pm} \prod_{i\in K^+} d_i^{+ d_i^{+(s)}} \prod_{i\in K^-} d_i^{- d_i^{-(s)}} \frac{1}{\prod_{j=0}^{k^+k^- - k^\pm + 1} s-j}\left(\frac{s!}{m!}\right)^2$$

$$\times \binom{m-s}{d_{i_1}^{-(s)},\ldots,d_{i_{k^-}}^{-(s)}}\binom{m-s}{d_{j_1}^{+(s)},\ldots,d_{j_{k^+}}^{+(s)}}.$$

It remains to bound $\frac{s!}{m!}$ and $\frac{\prod_{j=0}^{k^+k^--k^\pm+1} s-j}{m^{k^+k^--k^\pm}}$. First, using that $m-s=\mathcal{O}\left(d_{\max}^2\right)$, we find:

$$\begin{aligned}
\frac{m!}{s!} &= (s+1)(s+2)\cdots(m-1)m = m^{m-s}\left(1-\frac{1}{m}\right)\left(1-\frac{2}{m}\right)\cdots\left(1-\frac{m-s-1}{m}\right)\\
&= m^{m-s}\left(1-\prod_{i=1}^{m-s-1}\frac{i}{m}+\mathcal{O}\left((m-s)^2\frac{(m-s)^2}{m^2}\right)\right)\\
&\geq m^{m-s}e^{-\sum_{i=1}^{m-s-1}\frac{i}{m}+\mathcal{O}\left(\frac{d_{\max}^8}{m^2}\right)} = m^{m-s}e^{-\frac{(m-s)(m-s-1)}{2m}+\mathcal{O}\left(\frac{d_{\max}^8}{m^2}\right)} = m^{m-s}e^{-\mathcal{O}\left(\frac{d_{\max}^4}{m}\right)},
\end{aligned}$$

and therefore $\frac{s!}{m!} \leq \frac{1}{m^{m-s}}e^{\mathcal{O}\left(\frac{d_{\max}^4}{m}\right)} = \frac{1}{m^{m-s}}e^{o(1)}$. Second, let us consider $\frac{1}{\prod_{j=0}^{k^+k^--k^\pm+1} s-j}$. Using that $m-s\leq d_{\max}^2, k^+, k^-\leq d_{\max}$ and $0\leq k^\pm\leq \min\left(k^-,k^+\right)$, we obtain:

$$\begin{aligned}
\prod_{j=0}^{k^+k^--k^\pm+1} s-j \geq \prod_{j=0}^{k^+k^--k^\pm+1} m-d_{\max}^2-j &= m^{k^+k^--k^\pm}\prod_{j=0}^{k^+k^--k^\pm+1}\left(1-\frac{d_{\max}^2+j}{m}\right)\\
&= m^{k^+k^--k^\pm}\left(1-\prod_{j=1}^{k^+k^--k^\pm+1}\frac{d_{\max}^2+j}{m}+\mathcal{O}\left(\frac{d_{\max}^8}{m^2}\right)\right)\\
&\geq m^{k^+k^--k^\pm}e^{-\frac{(d_{\max}^2+k^+k^-+k^\pm+1)(d_{\max}^2+k^+k^-+k^\pm+2)}{2m}+\mathcal{O}\left(\frac{d_{\max}^8}{m^2}\right)}\\
&= m^{k^+k^--k^\pm}e^{-\mathcal{O}\left(\frac{d_{\max}^4}{m}\right)},
\end{aligned}$$

which gives

$$\frac{1}{\prod_{j=0}^{k^+k^--k^\pm+1} s-j} \leq \frac{1}{m^{k^+k^--k^\pm}}e^{\mathcal{O}\left(\frac{d_{\max}^4}{m}\right)} = \frac{1}{m^{k^+k^--k^\pm}}e^{o(1)}.$$

Thus the upper bound on the probability of $A_{d_{i_1}^{-\,(s)},\ldots,d_{i_{k^-}}^{-\,(s)},d_{j_1}^{+\,(s)},\ldots d_{j_{k^+}}^{+\,(s)}}$ becomes

$$e^{o(1)}d_{\max}^{2k^+k^--2k^\pm}\frac{\prod_{i\in K^+}d_i^{+\,d_i^{+\,(s)}}\prod_{i\in K^-}d_i^{-\,d_i^{-\,(s)}}}{m^{k^+k^--k^\pm}m^{2(m-s)}}\binom{m-s}{d_{i_1}^{-\,(s)},\ldots,d_{i_{k^-}}^{-\,(s)}}\binom{m-s}{d_{j_1}^{+\,(s)},\ldots,d_{j_{k^+}}^{+\,(s)}}.$$

□

Combining equation (49) with Lemma 3.2, we are able to prove the desired result.

Lemma 3.3. *The probability that Algorithm 1 returns a failure is* $o(1)$.

Proof. In the statement of Lemma 3.2, the fraction $\left(\frac{d_{\max}^2}{m}\right)^{k^+k^--k^\pm}$ is either 1 if $k^+k^-=k^\pm$ or smaller than $\frac{d_{\max}^2}{m}$ if $k^+k^-\neq k^\pm$. Since $k^\pm\leq\min\left(k^-,k^+\right)$, $k^+k^-=k^\pm$ implies that $k^+=k^-=1$. Together $k^+=k^-=1$ and the conditions under which the algorithm can fail imply that $K^+ = K^-$. First we consider this case. Since $K^+ = K^- = K^\pm = 1$ there holds ${d_{i_1}^-}^{(s)} = {d_{i_1}^+}^{(s)} = m - s$, plugging this into equation (50) we find

$$\mathbb{P}\left[A_{{d_{i_1}^-}^{(s)},{d_{i_1}^+}^{(s)}}\right] \leq e^{o(1)} \frac{{d_{i_1}^+}^{m-s} {d_{i_1}^-}^{m-s}}{m^{m-s} m^{m-s}} = o(1).$$

Next, assume that $k^+k^- \neq k^\pm$, which implies that $\left(\frac{d_{\max}^2}{m}\right)^{k^+k^- - k^\pm} \leq \frac{d_{\max}^2}{m}$. We apply the multinomial theorem to obtain:

$$\sum_{k^-=1}^{\max(m-s,d_{\max})} \sum_{i_1,\ldots,i_{k^-}=1}^{n} \prod_{i \in K^-} {d_i^-}^{{d_i^-}^{(s)}} \binom{m-s}{{d_{i_1}^-}^{(s)}, \ldots, {d_{i_{k^-}}^-}^{(s)}} = \left(d_1^- + \ldots + d_n^-\right)^{m-s}$$

and

$$\sum_{k^+=1}^{\max(m-s,d_{\max})} \sum_{j_1,\ldots,j_{k^+}=1}^{n} \prod_{i \in K^+} {d_i^+}^{{d_i^+}^{(s)}} \binom{m-s}{{d_{j_1}^+}^{(s)}, \ldots, {d_{j_{k^+}}^+}^{(s)}} = \left(d_1^+ + \ldots + d_n^+\right)^{m-s}.$$

Plugging these into equation (49) yields

$$\mathbb{P}\left[\,\text{failure}\,\right] \leq o(1) + e^{o(1)} \frac{d_{\max}^2}{m} \sum_{m-s=1}^{d_{\max}^2} \frac{\left(d_1^+ + \ldots d_n^+\right)^{m-s} \left(d_1^- + \ldots d_n^-\right)^{m-s}}{m^{m-s} m^{m-s}} \leq o(1).$$

□

This proves the claim of Theorem 1.1 about the failure probability of Algorithm 1.

4 Running time Algorithm 1

When implementing Algorithm 1 one has a certain freedom to chose how exactly choosing random samples with probability proportional to $P_{i,j}$ is performed. Our implementation of Algorithm 1 is based on the implementation of Bayati, Kim and Saberi [15] for undirected graphs, which, in turn, is based on Steger and Wormald's implementation for undirected regular random graphs. The latter uses a three-phase procedure, which depending on the step r, picks an edge (i, j) with probability proportional to ${d_i}^{(r)}{d_j}^{(r)}$ in a different manner. We also distinguish three phases depending on the algorithm step r, however, our sampling probability is proportional ${d_i^+}^{(r)}{d_j^-}^{(r)}\left(1 - \frac{d_i^+ d_j^-}{2m}\right)$, and the corresponding criteria that determine the phase of the algorithm are different. In what follows, we show that the expected running time of our algorithm is $\mathcal{O}\left(md_{\max}\right)$, that is we prove the following lemma.

Lemma 4.1. *Algorithm 1 can be implemented so that its expected running time is* $\mathcal{O}\left(md_{\max}\right)$ *for graphical degree sequences* $\mathbf{d}$ *with* $d_{\max} = \mathcal{O}\left(m^{1/4-\tau}\right)$ *for some* $\tau > 0$.

Proof. Phase 1. Let E be the list of edges constructed by the algorithm so far. In the first phase, a random unmatched in- and out-stubs are selected. We may check whether this is an eligible pair in time $\mathcal{O}\left(d_{\max}\right)$. If eligible, the pair is accepted with probability proportional to $1-\frac{d_i^+ d_j^-}{2m}$ and (i,j) is added to E. We select edges according to this procedure until the number of unmatched in-stubs drops bellow $2d_{\max}^2$. This marks the end of phase 1. As a crude estimate, each eligible pair is accepted with probability at least $\frac{1}{2}$, and at most $\frac{1}{2}$ of all stub pairs is ineligible, see Lemma 2.3(a). Hence, creating one edge in phase 1 has an expected computational complexity of $\mathcal{O}(d_{\max})$, and the total runtime of this phase is $\mathcal{O}(md_{\max})$.

Phase 2. In this phase we select a pair of vertices instead of a pair of stubs. This requires us to keep track of the list of vertices with unmatched in-stubs/out-stubs. These lists are constructed in $\mathcal{O}(n)$ and can be updated in a constant time. Draw uniformly random vertices i and j from the lists of vertices with unmatched out-stubs and in-stubs correspondingly. Accept i (respectively j) with probability $\frac{d_i^{+(r)}}{d_i^{+(r)}}\left(\frac{d_j^{-(r)}}{d_j^{-(r)}}\right)$. If both vertices are accepted, we check if (i,j) is an eligible edge in time $\mathcal{O}(d_{\max})$. If the edge is eligible, it is accepted with probability $1-\frac{d_i^+ d_j^-}{2m}$. Phase 2 ends when the number of vertices with unmatched in-stubs or the number of vertices with unmatched out-stubs is less than $2d_{\max}$. Since every vertex with unmatched in-stubs (respectively out-stubs) has at most $d_{\max}$ unmatched in-stubs (out-stubs), this guarantees that the edge is eligible with probability at least $\frac{1}{2}$. To get a pair of accepted vertices, we need an expected number of $\mathcal{O}\left(d_{\max}^2\right)$ redraws. Thus the construction of one edge is expected to take $\mathcal{O}\left(d_{\max}^2\right)$. As there are only $2d_{\max}^2$ unmatched in-stubs at the start of phase 2, at most $d_{\max}^2$ edges are created in this phase. Thus the expected running time of Phase 2 is $\mathcal{O}\left(d_{\max}^4\right)$.

Phase 3. At the beginning of this phase, a list $\widetilde{E}$ of all remaining eligible edges is constructed. At the start of phase 3 there are only $2d_{\max}$ vertices left with unmatched in-stubs or with unmatched out-stubs. Hence there are at most $2d_{\max}^2$ vertices with unmatched out-stubs or in-stubs. Thus $\widetilde{E}$ contains no more than $4d_{\max}^3$ edges. For each possible edge we check in time $\mathcal{O}(d_{\max})$ if it does not create a double edge or self-loop. Thus, constructing $\widetilde{E}$ takes $\mathcal{O}\left(d_{\max}^4\right)$. The rest of Phase 3 consist of picking a random element of $\widetilde{E}$ and accepting it with probability $\frac{d_i^{+(r)} d_j^{-(r)}}{d_i^+ d_j^-}\left(1-\frac{d_i^+ d_j^-}{2m}\right)$. This leads to an expected number of $\mathcal{O}\left(d_{\max}^2\right)$ repetitions to accept one edge. If an edge is accepted, it is removed from $\widetilde{E}$ and the values of $d_i^{+(r)}$ and $d_j^{-(r)}$ are updated. After selecting an element of $\widetilde{E}$, it must be checked if $d_i^{+(r)}>0$ and $d_j^{-(r)}>0$. If this is not the case, the edge is not added to E and removed from $\widetilde{E}$. This continues until $\widetilde{E}$ is empty or $|E|=m$. This has expected running time of order $\mathcal{O}\left(d_{\max}^5\right)$ as there are $\mathcal{O}\left(d_{\max}^3\right)$ edges that are expected to be discarded or accepted in $\mathcal{O}\left(d_{\max}^2\right)$. Thus, the total running time of the algorithm is $\mathcal{O}(md_{\max})+\mathcal{O}(n)+\mathcal{O}\left(d_{\max}^4\right)+\mathcal{O}\left(d_{\max}^5\right)$. As $d_{\max}=\mathcal{O}\left(m^{1/4-\tau}\right)$, the running time is $\mathcal{O}(md_{\max})$.

We must also compute P_{ij} at each step. Let $P_{ij}^{(r)}$ denote the probabilities that the edge (i,j) is added to E at step r. There holds:

$$P_{ij}^{(r)}=\frac{d_i^{+(r)} d_j^{-(r)}\left(1-\frac{d_i^+ d_j^-}{2m}\right)}{(m-r)^2-\Psi_r(\mathcal{N})}. \tag{52}$$

The numerator $d_i^{+(r)} d_j^{-(r)}\left(1-\frac{d_i^+ d_j^-}{2m}\right)$ can be computed in a constant time. To determine the denominator in (52), remark that:

$$\begin{aligned}&\left[(m-r+1)^2-\Psi_{r+1}\left(\mathcal{N}\right)\right]-\left[(m-r)^2-\Psi_r\left(\mathcal{N}\right)\right]\\&=\sum_{(u,v)\in E_{r+1}} d_u^{+(r+1)}d_v^{-(r+1)}\left(1-\frac{d_u^+d_v^-}{2m}\right)-\sum_{(u,v)\in E_r} d_u^{+(r)}d_v^{-(r)}\left(1-\frac{d_u^+d_v^-}{2m}\right)\\&+\sum_{(i,v)\in G_{\mathcal{N}_r}} d_v^{-(r)}\left(1-\frac{d_i^+d_v^-}{2m}\right)+\sum_{(u,j)\in G_{\mathcal{N}_r}} d_u^{+(r)}\left(1-\frac{d_u^+d_j^-}{2m}\right)+d_i^{-(r)}\left(1-\frac{d_i^+d_j^-}{2m}\right)\\&+d_j^{+(r)}\left(1-\frac{d_i^+d_j^-}{2m}\right).\end{aligned}$$

At each step r, each of the terms in the latter expression can be updated in $\mathcal{O}\left(d_{\max}\right)$ operations. This allows us to determine the value of $P_{ij}^{(r)}$ in time $\mathcal{O}\left(d_{\max}\right)$. As the construction of one edge also takes at least $\mathcal{O}\left(d_{\max}\right)$ in every phase, this does not change the overall complexity of the algorithm. The initial value is

$$\Psi_0\left(\mathcal{N}\right)=m^2-\sum_{i=1}^n d_i^- d_i^+ - \frac{\sum_{i=1}^n {d_i^-}^2\sum_{i=1}^n {d_i^+}^2-\sum_{i=1}^n {d_i^-}^2{d_i^+}^2}{2m},$$

which can be computed in $\mathcal{O}\left(n\right)$. As $n\le m$ this does not change the order of the expected running time, and hence, this completes the proof. □

This lemma completes the proof of Theorem 1.1.

Acknowledgements

We are grateful to Péter L. Erdős for an interesting discussion about MCMC algorithms.

References

[1] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.

[2] Brendan D McKay and Nicholas C Wormald. Uniform generation of random regular graphs of moderate degree. *Journal of Algorithms*, 11(1):52–67, 1990.

[3] Pu Gao and Nicholas Wormald. Uniform generation of random regular graphs. *SIAM Journal on Computing*, 46(4):1395–1427, 2017.

[4] Pu Gao and Nicholas Wormald. Uniform generation of random graphs with power-law degree sequences. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1741–1758. SIAM, 2018.

[5] Ching Law and K-Y Siu. Distributed construction of random expander networks. In *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No. 03CH37428)*, volume 3, pages 2133–2143. IEEE, 2003.

[6] Prateek Jain, Praneeth Netrapalli, and Sujay Sanghavi. Low-rank matrix completion using alternating minimization. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 665–674, 2013.

[7] Jürgen Lerner. Role assignments. In *Network analysis*, pages 216–252. Springer, 2005.

[8] Fernando L Metz, Giorgio Parisi, and Luca Leuzzi. Finite-size corrections to the spectrum of regular random graphs: An analytical solution. *Physical Review E*, 90(5):052109, 2014.

[9] Tim Rogers, Conrad Pérez Vicente, Koujin Takeda, and Isaac Pérez Castillo. Spectral density of random graphs with topological constraints. *Journal of Physics A: Mathematical and Theoretical*, 43(19):195002, 2010.

[10] Jacopo Grilli, Tim Rogers, and Stefano Allesina. Modularity and stability in ecological communities. *Nature communications*, 7(1):1–10, 2016.

[11] Gottfried Tinhofer. On the generation of random graphs with given properties and known distribution. *Appl. Comput. Sci., Ber. Prakt. Inf*, 13:265–297, 1979.

[12] A Ramachandra Rao, Rabindranath Jana, and Suraj Bandyopadhyay. A markov chain monte carlo method for generating random (0, 1)-matrices with given marginals. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 225–242, 1996.

[13] Annabell Berger and Matthias Müller-Hannemann. Uniform sampling of digraphs with a fixed degree sequence. In *International Workshop on Graph-Theoretic Concepts in Computer Science*, pages 220–231. Springer, 2010.

[14] Angelika Steger and Nicholas C Wormald. Generating random regular graphs quickly. *Combinatorics, Probability and Computing*, 8(4):377–396, 1999.

[15] Mohsen Bayati, Jeong Han Kim, and Amin Saberi. A sequential algorithm for generating random graphs. *Algorithmica*, 58(4):860–910, 2010.

[16] Andrii Arman, Pu Gao, and Nicholas Wormald. Fast uniform generation of random graphs with given degree sequences. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1371–1379. IEEE, 2019.

[17] Tinhofer Gottfried. On the generation of random graphs with given properties and known distribution. *Appl. Comput. Sci., Ber. Prakt. Inf.*, 13:265–297, 1979.

[18] Mark Jerrum and Alistair Sinclair. Fast uniform generation of regular graphs. *Theoretical Computer Science*, 73(1):91–100, 1990.

[19] Colin Cooper, Martin Dyer, and Catherine Greenhill. Sampling regular graphs and a peer-to-peer network. *Comb. Probab. Comput.*, 16(4):557–593, 2007.

[20] Péter L Erdös, Istán Miklós, and Lajos Soukup. Towards random uniform sampling of bipartite graphs with given degree sequence. *Electronic Journal of Combinatorics*, 20:P16, 2013.

[21] Ravi Kannan, Prasad Tetali, and Santosh Vempala. Simple markov-chain algorithms for generating bipartite graphs and tournaments. *Random Structures & Algorithms*, 14(4):293–308, 1999.

[22] Péter L Erdős, Catherine Greenhill, Tamás Róbert Mezei, István Miklós, Dániel Soltész, and Lajos Soukup. The mixing time of the switch markov chains: a unified approach. *arXiv preprint arXiv:1903.06600*, 2019.

[23] Mark Jerrum, Brendan D McKay, and Alistair Sinclair. *When is a graphical sequence stable?* University of Edinburgh, Department of Computer Science, 1989.

[24] Pu Gao and Catherine Greenhill. Mixing time of the switch markov chain and stable degree sequences. *Discrete Applied Mathematics*, 291:143–162, 2021.

[25] Svante Janson. Random graphs with given vertex degrees and switchings. *Random Structures & Algorithms*, 57(1):3–31, 2020.

[26] Catherine Greenhill. A polynomial bound on the mixing time of a markov chain for sampling regular directed graphs. *Electronic Journal of Combinatorics*, 18:P234, 2011.

[27] Catherine Greenhill and Matteo Sfragara. The switch markov chain for sampling irregular graphs and digraphs. *Theoretical Computer Science*, 719:1–20, 2018.

[28] Péter L. Erdős, Tamás Róbert Mezei, István Miklós, and Dániel Soltész. Efficiently sampling the realizations of bounded, irregular degree sequences of bipartite and directed graphs. *PLOS ONE*, 13(8):1–20, 08 2018.

[29] Éva Czabarka, Aaron Dutle, Péter L Erdős, and István Miklós. On realizations of a joint degree matrix. *Discrete Applied Mathematics*, 181:283–288, 2015.

[30] Georgios Amanatidis, Bradley Green, and Milena Mihail. Graphic realizations of joint-degree matrices. *arXiv preprint arXiv:1509.07076*, 2015.

[31] Charo I. Del Genio, Hyunju Kim, Zoltán Toroczkai, and Kevin E. Bassler.

[32] Kevin E Bassler, Charo I Del Genio, Péter L Erdős, István Miklós, and Zoltán Toroczkai. Exact sampling of graphs with prescribed degree correlations. *New Journal of Physics*, 17(8):083052, 2015.

[33] Joseph Blitzstein and Persi Diaconis. A sequential importance sampling algorithm for generating random graphs with prescribed degrees. *Internet mathematics*, 6(4):489–522, 2011.

[34] Jeong Han Kim and Van H Vu. Generating random regular graphs. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 213–222, 2003.

[35] M Drew LaMar. Directed 3-cycle anchored digraphs and their application in the uniform sampling of realizations from a fixed degree sequence. In *Proceedings of the 2011 Winter Simulation Conference (WSC)*, pages 3348–3359. IEEE, 2011.

[36] Van H Vu. Concentration of non-lipschitz functions and applications. *Random Structures & Algorithms*, 20(3):262–316, 2002.