『経済研究』（明治学院大学）第 160 号　2020 年

特定の次数列を満たすランダムグラフの大きさについて

浜　口　幸　弘

1. はじめに

本稿の目的は以前の論文（浜口（2014），（2016））を修正改善し，さらに新たな知見を提案することにある．本稿では，コンフィギュレーションモデルによって構成されるランダムグラフの大きさに関する定理 1 および定理 2 を導出し，さらに定理 1 に基づきランダムグラフの大きさに関する予想を示している．

2. コンフィギュレーションモデル

次数列が与えられた頂点数 n のランダムグラフ G の構成は，一般にコンフィギュレーションモデルによって実現される．n 頂点および次数列 $\boldsymbol{D}$ のランダムなコンフィギュレーション F は次のように定義される．

頂点 v に対して，その次数を $\deg(v)$ で表す．このとき，$\deg(v)$ 個のコピーを作り，全体のコピーの集合を L とする．そして, L の 2 つのコピーをランダムに選んでマッチング対を構成する．すべてのマッチングがなされたとき，コンフィギュレーション F が完成する．よって，次数の総和は偶数となる必要があるので，以降，この条件は満たされているものとする．また，1 つのマッチングの対がグラフにおける 1 本の辺に対応するので，得られる各 F は多重グラフを表している．すなわち，1 頂点におけるループや 2 つの頂点を結ぶ 2 本以上の辺を認めている．しかし，一定条件下において，ランダムな多重グラフのある性質 R を対象にすることで，ランダムな単純グラフの性質 R を議論できるということは，よく知られた事実である（例えば，Frize and Karonski（2016）を参照）．なお，R に対してその確率が $\lim_{n\to\infty} P(R)=1$ のとき，本稿では「ほとんどすべてのコンフィギュレーション（またはグラフ）において，R が成り立つ」と表現する．

そこで，Molloy and Reed（2000）による F の構成アルゴリズムを説明する．まず，次のように各用語を定義する．頂点 v に対して，そのすべてのコピーがマッチングされているならば，v は“completely exposed（完全開示）”状態にあるといい，コピーのすべてではなく一部がマッチングされているならば，v は“partially exposed（部分開示）”状態にあるという．よって，それ以外の頂点は“unexposed（未開示）”状態にある．また，部分開示の頂点のコピーにおいて，まだマッチングされていないコピーを“open”状態にあるという意味でオープンコピーと呼ぶことにする．そして，F を構成するアルゴリズムは以下のようになる．

1. 各頂点 v に対して，その次数分 $\deg(v)$ 個のコピーを作り，全体のコピーの集合を L とする．
2. L の要素が尽きるまで以下の手続きを繰り返す．

(a) L の任意の要素 1 つを選択し，次に，その対となるもう 1 つの要素を L からランダムに選択する．
　これら 2 つの要素（マッチング対）を L から除外して(b)に行く．

(b) 部分開示の頂点がある限り以下の手続きを繰り返す．尽きれば(a)に戻る．
　部分開示にある頂点のオープンコピーを 1 つ任意に選択し（ランダムでなくてもよい），その対となる要素を L からランダムに選択する．これら 2 つの要素（マッチング対）を L から除外する．

このアルゴリズムでは，最初に L の 1 つのコピーが任意に選択され，次にその対となるコピーが L からランダムに選択される手続きが 1 つの試行（1 ステップ）をなす．そして，1 つの連結成分が完成してから，新たな次の連結成分の構成に取り掛かることになる（手続き 2(a)に戻る）．また，各頂点の 2 つのコピー間で起こり得るマッチングはすべて等確率であり，各ステップにおける任意のコピーの選択では，より先に選択された頂点におけるオープンコピーを優先的に選択するものとする．

F の構成アルゴリズムにおいて，t 番目のコピー対が決定された直後のオープンコピーの個数を X_t とし，ここまで構成された部分的コンフィギュレーションを F_t と記す．このとき，手続き 2(b)において 2 番目にランダム選択されるコピーの頂点（次数 d とする）が未開示ならば，X_t は $d-2$ 増加し，部分開示ならば，X_t は 2 減少する．したがって，アルゴリズムを通してつねに $X_t \geq 0$ である．また，後者の頂点コピーの対を“backedge”と呼ぶ．バックエッジが作られて，オープンコピーがなくなるときに，1 つの連結成分が生成されることになる．

3. Molloy and Reed の定理の概括

ここでは，本稿で考察の対象とする Molloy and Reed（2000）による定理を概括する．次数列が与えられた頂点数 n のランダムグラフ G に対して，次数 i の頂点の個数を $d_i(n)$ と記し，その次数列を $\boldsymbol{D}= d_0(n), d_1(n), \cdots, d_i(n), \cdots$ とする（$i \leq n-1$）．一方，頂点 v の次数は $\deg(v)$ と記す．また，$n=\sum_i d_i(n)$ は十分に大きな数とする．そして Molloy and Reed（2000）は次数列 $\boldsymbol{D}$ の条件を次のように与えている．

$\boldsymbol{D}$ に対して，$\lim_{n\to\infty} d_i(n)/n=\lambda_i$ となるような定数 $\lambda_i \geq 0$ が存在し，さらに，$\sum_{i\geq 1} id_i(n)/n=K+o(1)$ となる正の定数 K が存在する．また，次数列 $\boldsymbol{D}$ は“well-behaved”（Molloy and Reed（2000））である．

次に，G に大規模な連結成分(component)が含まれることを評価する基準量として，$Q(\boldsymbol{D})=\sum_{i\geq 1} i(i-2)\lambda_i$ を設定している．このとき，$Q(\boldsymbol{D})>0$ ならば，$\lambda_i=\lim_{n\to\infty} d_i(n)/n>0$ すなわち，ある正の定数 c に対して，$d_i(n)/n\to c$ となる $i\geq 3$ が少なくとも 1 つ存在する．また，$\lim_{n\to\infty} d_i(n)/n=0$ となる i は $Q(\boldsymbol{D})$ の大きさに影響しない．すなわち，この条件を満たす頂点の存在は，その有無に関わらず Molloy and Reed（2000）の定理の内容には影響しない．その定理は以下のようになる．

定理　Molloy and Reed（2000）

次数列 $\boldsymbol{D}$ は，任意の n およびある $\varepsilon>0$ に対して，$i>n^{1/4-\varepsilon}$ のとき $d_i(n)=0$ となるような，$\boldsymbol{D}$ の条件を満たす次数列とする．また，G は頂点数 n で次数 i の頂点を $d_i(n)$ 個もつランダムグラフとする．このとき，

(a) $Q(\boldsymbol{D})>0$ ならば，ある定数 $\zeta_1>0$ および $\zeta_2>0$ に対して，ほとんどすべての G は少なくとも $\zeta_1 n$ 個の頂点および $\zeta_2 n$ 個の閉路をもつ連結成分を含む．また，$Q(\boldsymbol{D})$ が有限ならば，ほとんどすべての G は，ある定数 $\gamma>0$ に対して，$\gamma \log n$ より大きい連結成分をただ 1 つもつ．

(b) $Q(\boldsymbol{D})<0$ かつ，ある $0\leq\omega(n)\leq n^{1/8-\varepsilon}$ に対して，$d_i(n)=0$ $(i\geq\omega(n))$ ならば，ある定数 $R>0$ に対して，ほとんどすべての G は $R\omega(n)^2\log n$ 個以上の頂点をもつ連結成分をもたない，また，ほとんどすべての G は $2R\omega(n)^2\log n$ 個以上の閉路をもたない．また，ほとんどすべての G において，どの連結成分も高々1 個の閉路しかもたない．

4. 次数 1 と次数 k の頂点からなるランダムグラフに関する考察

本稿では前提条件を簡単化して，ランダムグラフが次数 1 と次数 k の頂点から構成される場合を考える．次数 1 の頂点はコピーを 1 つ持つだけなので，そのコピーが 1 回選択されたら頂点は完全開示になり，次にその頂点が選択されることはない．よって，次数 1 の頂点の占める割合が連結成分の形成に大きな影響を及ぼすことになる．これは Molloy and Reed（2000）の場合と同じである．また，次数 k は単なる定数ではなく，n の関数とする．

さて以下では，頂点数 n のランダムグラフ G が，$d_1(n)$ 個の次数 1 の頂点および $d_k(n)$ 個の次数 k の頂点から構成される場合について考察を行う．まず，M を最初の状態における総コピー数とし，次数 k の値は，$4\leq k\leq n^{1/8}$ を満たすものとする（$k=3$ の場合は，補題 2 の証明に用いる定理において，その前提条件を満たさないので除外する）．次に，基準値 α を以下のように定義する．これは $Q(\boldsymbol{D})$ の定義に似ているが，本稿独自のものである．

$\alpha=\sum_{i=1}^{n-1} i(i-2)d_i(n)/n=-d_1(n)/n+k(k-2)d_k(n)/n$ として $\lim_{n\to\infty}\alpha$ は，$0<\lim_{n\to\infty}\alpha<1/2$ を満たすように収束する（なお以下では，十分大きな n を前提にするので，$0<\alpha<1/2$ と考える）．このとき，

$d_1(n)+d_k(n)=n$ かつ $M=d_1(n)+kd_k(n)$ なので，以下を得る．

$$d_1(n)=\frac{k(k-2)-\alpha}{(k-1)^2}n,\quad d_k(n)=\frac{\alpha+1}{(k-1)^2}n,\quad M=\frac{k+\alpha}{k-1}n.$$

また，$n < M \leq 3n/2$ である．$0<\alpha<1/2$ とする理由は，Molloy and Reed（2000）の結果から判断して，おおよそこの条件を満たす α が頂点数の大きな連結成分の生成される境界付近の値だからである．

ここで，ランダムな多重グラフの性質 R を対象にすることで，ランダムな単純グラフの性質 R を議論できることを，Frize and Karonski（2016）の定理 10.3 に従い確認すると，以下のようになる．ここでは頂点 i の次数を d_i と表すことに注意．

$$M_1=\sum_i d_i=\frac{k(k-2)-\alpha}{(k-1)^2}n+\frac{(\alpha+1)k}{(k-1)^2}n=\frac{k+\alpha}{k-1}n.$$

$$M_2=\sum_i d_i(d_i-1)=k(k-1)\frac{\alpha+1}{(k-1)^2}n=\frac{(\alpha+1)k}{k-1}n.$$

よって，$\lambda=M_2/(2M_1)=O(1)$ となり，その条件を満たしている．

さて本稿のモデルと Molloy and Reed（2000）との相違点をまとめると以下である．すなわち，本稿では，

1. ランダムグラフ G は，その次数列を簡略化して次数 1 と次数 k の頂点のみから構成される．

α に負の影響を与える次数 1 の頂点の存在は同じで，正の影響を与える頂点は次数 k の頂点のみとする．このとき，α に対する両方の影響要因を扱っているので，一般性はそれほど失われないと考える．

2. ある次数 i に対して，$\lim_{n\to\infty} d_i(n)/n=0$ かつ $\lim_{n\to\infty} i(i-2)d_i(n)/n>0$ となる場合も議論に含める．

すなわち，次数列 $\boldsymbol{D}$ は“well-behaved”（Molloy and Reed（2000））という定義における 2 番目の条件を除外する．例えば，$d_{\sqrt{\log n}}(n)=n/\log n$ の場合である．この場合は，Molloy and Reed（2000）における確率的優位性（stochastic dominance）を用いた証明方法では対応できないので，本稿の特徴の 1 つとなる．

さて，第 t $(t\geq 0)$ 回目の試行の結果，部分開示の頂点を含む連結成分（高々，1 つ存在）に含まれるオープンコピーの総数を X_t と定義した．このとき，$X_i>0$ $(0\leq i\leq t-1)$ の場合に対して，前述のアルゴリズムは以下のように関数 $\mu(x)$ を用いて定式化できる．ただし，その手続き 2. (a) における最初の要素の任意選択では，次数 k の未開示な頂点のコピーを優先的に選択するものとする．

$$\mu(x)=\begin{cases} k-2 & : L \text{ から次数 } k \text{ の未開示な頂点のコピー } x \text{ を選択する場合,} \\ -1 & : L \text{ から次数 } 1 \text{ の未開示な頂点のコピー } x \text{ を選択する場合,} \\ -2 & : L \text{ から部分開示にある頂点のコピー } x \text{ を選択する場合,} \end{cases}$$

$$X_0=k,\quad X_t=X_{t-1}+\mu(x)\quad (x\in L).$$

ただし，X_t が 0 になった時点でその連結成分の生成は完了し，改めて，$X_t=k$ と設定し直して，第 $(t+1)$ 回目のコピーの選択から新しい連結成分の生成が開始される．

さらに，2 つの確率変数を定義すると，Y_t は第 t 回目の試行後の次数 1 の未開示な頂点におけるコピーの総数を表し，Z_t は第 t 回目の試行後の次数 k の未開示な頂点におけるコピーの総数を表す．次に，これらの変数に関連する事象を定義する．A_t は第 t 回目の試行においてオープンコピーがランダムに選択

される事象であり，B_t は第 t 回目の試行において次数 1 の未開示な頂点のコピーがランダムに選択される事象であり，C_t は第 t 回目の試行において次数 k の未開示な頂点のコピーがランダムに選択される事象である．これらの定義に基づいて以下のような式を構成できる．

まず，本稿を通して，前述のアルゴリズムの手続き 2．(a)における L の最初の要素を任意に選択するときは，次数 k の未開示な頂点のコピーを優先的に選択するものとする．そして，各ステップにおける任意のコピーの選択では，より先に選択された頂点におけるオープンコピーを優先的に選択するものとする．また，単に「各ステップにおける頂点の選択」と表す場合は，そのステップにおける 2 番目のコピーのランダムな選択を意味するものとする．このとき，以下の式が成り立つ．

$$\begin{aligned}
&M-2t=X_t+Y_t+Z_t,\\
&X_{t+1}=X_t-2I_{A_{t+1}}-I_{B_{t+1}}+(k-2)I_{C_{t+1}}+kI_{\{X_t=0\}},\\
&Y_{t+1}=Y_t-I_{B_{t+1}},\\
&Z_{t+1}=Z_t-kI_{C_{t+1}}-kI_{\{X_t=0\}}.
\end{aligned} \tag{1}$$

ただし，I_E は事象 E が起こるとき 1 をとり，そうでないとき 0 をとる指示関数を表す．

また，$X_t>0$ のとき，以下のような条件付確率に関する式が成り立つ．ただし，F_t は t 番目のコピー対が決定された直後のコンフィギュレーションとする．

$$\begin{aligned}
P(A_{t+1}|F_t)&=\frac{X_t-1}{M-2t-1},\\
P(B_{t+1}|F_t)&=\frac{Y_t}{M-2t-1},\\
P(C_{t+1}|F_t)&=\frac{Z_t}{M-2t-1}.
\end{aligned} \tag{2}$$

このとき，バックエッジの有無に関する以下の補題 1 を得る．

補題 1

$0<\delta<\dfrac{1}{8}$ に対して，σ を $\sigma=n^{\frac{9}{16}+\delta}$ とする．試行回数 t が $1\leq t\leq M/\sigma$ の場合，ほとんどすべてのコンフィギュレーション F_t において，2 個のオープンコピーのマッチングは行われない．

証明

F の構成アルゴリズムにおいて，第 t ステップまでバックエッジのない確率を $P_{NB}(t)$ とする．前述の頂点コピー選択の定義から，最初に次数 k の頂点のコピーを選択する．このとき，$P_{NB}(t)$ の下限を評価するには，各ステップにおいてオープンコピーが最大になる場合を考え，その最大数を選択可能なコピー数から引いてバックエッジを選択しない確率を求め，かけていけばよい．よって，以下の評価を得る．

$$P_{NB}(t)\geq\prod_{i=1}^{t}\frac{M-(2i-1)-(X_{i-1}-1)}{M-(2i-1)}$$

$$\geq \frac{M-1-(k-1)}{M-1}\ \frac{M-3-(k-1+(k-2))}{M-3}\cdots\frac{M-(2t-1)-(k-1+(t-1)(k-2))}{M-(2t-1)}$$

$$\geq\left(1-\frac{t(k-1)}{M-(2t-1)}\right)^t\geq 1-\frac{t^2(k-1)}{M-(2t-1)}\geq 1-2\frac{t^2k}{M}=1+O\left(\frac{kM}{\sigma^2}\right).$$

ここで，$kM\leq\dfrac{3}{2}n^{9/8}$ なので，$\dfrac{kM}{\sigma^2}\leq\dfrac{3}{2n^{2\delta}}$．よって，主張は成り立つ．□

補題 1 を用いて以下の補題 2 を導く．なお，これ以降，$t=M/\sigma$ のように等号で結ばれた表記において，左辺が整数であり，右辺が分数になる場合は，証明において問題にならない限り，そのままの表記で右辺も整数とみなす．

補題 2

ε を $0<\varepsilon<\min\left\{\dfrac{3\alpha}{4(\alpha+1)},\dfrac{1}{12}\right\}$ を満たす任意の定数とする．$\sigma=n^{\frac{9}{16}+\delta}\ \left(0<\delta<\dfrac{1}{8}\right)$，$t_0=\dfrac{M}{\sigma}$ および $h=\dfrac{9((\alpha-\varepsilon\alpha-\varepsilon)k-\alpha)}{10(k-1)}\ \dfrac{n}{\sigma}$ とすると，ほとんどすべてのコンフィギュレーション F_{t_0} において，$X_{t_0}>h$ が成り立つ．

証明

補題 1 から，$0\leq t\leq t_0$ において，ほとんどすべてのコンフィギュレーションはバックエッジを持たない．以下では，このことを前提条件とするので，扱う確率は条件付確率として表記すべきだが，表記を簡単にするために省略することにする．そこで，第 t_0 ステップまでに未開示の次数 k の頂点のコピーをランダムに選択する合計数を $R_k^{(t_0)}$ とし，まず，$R_k^{(t_0)}$ の下限を評価する．いま，$R_k^{(t_0)}$ を分解して，$R_k^{(t_0)}=\sum_{t=1}^{t_0}R_t$ とする．R_t は第 t 番目のステップにおいて次数 k の頂点のコピーを選択する場合は 1，そうでない場合は 0 をとる確率変数とする．$R_1,\cdots,R_{t_0}$ は独立でないことに注意する．次に，R_t に対応する確率変数 Q_t を次のように定める．

$$P(Q_t=1)=\frac{kd_k(n)-kt_0}{M},\quad P(Q_t=0)=1-P(Q_t=1).$$

そして，$Q_k^{(t_0)}=\sum_{t=1}^{t_0}Q_t$ として $Q_k^{(t_0)}$ を定める．このとき，$Q_1,\cdots,Q_{t_0}$ は互いに独立である．よって，以下が成り立つ．

$$P(Q_t=1)\leq P(R_t=1\,|\,R_1,\cdots,R_{t-1}).$$

すなわち，負でない実数 x に対して，

$$P(Q_t\geq x)\leq P(R_t\geq x\,|\,R_1,\cdots,R_{t-1}).$$

これは，確率的優位性（stochastic dominance）の条件（例えば，Frieze and Karonski（2016）を参照）を満たすので，以下を得る．

$$P(Q_k^{(t_0)} \geq x) \leq P(R_k^{(t_0)} \geq x).$$

よって，任意の$0 < \varepsilon < \min\left\{\frac{3\alpha}{4(\alpha+1)}, \frac{1}{12}\right\}$に対して，

$$P\left(R_k^{(t_0)} \geq (1-\varepsilon)E\left(Q_k^{(t_0)}\right)\right) \geq P\left(Q_k^{(t_0)} \geq (1-\varepsilon)E\left(Q_k^{(t_0)}\right)\right).$$

ここで，$E\left(Q_k^{(t_0)}\right) = t_0 \dfrac{kd_k(n) - kt_0}{M} = \dfrac{kd_k(n)}{\sigma}\left(1 + O\left(\dfrac{k^2}{\sigma}\right)\right)$である．

そして，$p = \dfrac{kd_k(n) - kt_0}{M}$とすると，$0 < \varepsilon < \min\left\{\dfrac{3\alpha}{4(\alpha+1)}, \dfrac{1}{12}\right\}$，$4 \leq k \leq n^{\frac{1}{8}}$および$0 < \alpha < 1/2$ なので，Bollobás（2001）にある定理 1.7 の条件を満たし，それを適用できる．よって，

$$P\left(Q_k^{(t_0)} \geq (1-\varepsilon)E\left(Q_k^{(t_0)}\right)\right) \geq 1 - \left(\frac{\varepsilon^2 n^{5/16-\delta}}{2}\right)^{-1/2} e^{-\varepsilon^2 n^{5/16-\delta}/6}.$$

ただし，$pt_0 \geq \dfrac{n^{5/16-\delta}}{2}$ と評価している．また，$k=3$ の場合，$p \leq 1/2$ の条件を満たさないことに注意.

したがって，

$$P\left(R_k^{(t_0)} \geq (1-\varepsilon)E\left(Q_k^{(t_0)}\right)\right) \geq 1 - \left(\frac{\varepsilon^2 n^{5/16-\delta}}{2}\right)^{-1/2} e^{-\varepsilon^2 n^{5/16-\delta}/6}$$

となるので，ほとんどすべての F_{t_0} において，以下が成り立つ.

$$R_k^{(t_0)} \geq (1-\varepsilon)\frac{kd_k(n)}{\sigma}\left(1 + O\left(\frac{k^2}{\sigma}\right)\right).$$

一方，$R_k^{(t_0)}$の上限を評価する場合，R_t に対応する確率変数 Q_t を改めて次のように定める.

$$P(Q_t = 1) = \frac{kd_k(n)}{M - 2t_0 + 1}, \quad P(Q_t = 0) = 1 - P(Q_t = 1).$$

以降は，上述と同じように考えれば，ほとんどすべての F_{t_0} において，以下が成り立つ.

$$R_k^{(t_0)} \leq (1+\varepsilon)\frac{kd_k(n)}{\sigma}\left(1 + O\left(\frac{1}{\sigma}\right)\right).$$

さて，上記のことから以下を得る.

$$P\left((k-1)R_k^{(t_0)} - t_0 \geq (1-\varepsilon)(k-1)E\left(Q_k^{(t_0)}\right) - t_0\right) \geq 1 - \left(\frac{\varepsilon^2 n^{5/16-\delta}}{2}\right)^{-1/2} e^{-\varepsilon^2 n^{5/16-\delta}/6}.$$

ところで，$0 \leq t \leq t_0$ において，X_t の大きさを評価すると，次数 k の頂点のコピーを 1 回選択するごとに少なくとも $k-2$ 個のオープンコピーが得られ（それ以前のオープンコピーが 0 個の場合は，$2(k-1)$個のオープンコピーが得られる)，次数 1 の頂点のコピーを 1 回選択すれば高々1 個のオープンコピーを失うので (それ以前のオープンコピーが 0 個の場合は，$k-2$ 個のオープンコピーが得られる)，以下を得る.

$$X_t \geq (k-2)R_k^{(t)} + (-1)\cdot(t - R_k^{(t)}) = (k-1)R_k^{(t)} - t.$$

これを上述の結果に適用すると，

$$P\left(X_{t_0}\geq (1-\varepsilon)(k-1)E\left(Q_k^{(t_0)}\right)-t_0\right)\geq 1-\left(\frac{\varepsilon^2 n^{5/16-\delta}}{2}\right)^{-1/2} e^{-\varepsilon^2 n^{5/16-\delta}/6}.$$

このとき，

$$\begin{aligned}(1-\varepsilon)(k-1)E\left(Q_k^{(t_0)}\right)-t_0 &= (1-\varepsilon)(k-1)\frac{kd_k(n)}{\sigma}\left(1+O\left(\frac{k^2}{\sigma}\right)\right)-\frac{d_1(n)+kd_k(n)}{\sigma}\\ &=\frac{(1-\varepsilon)k(k-1)d_k(n)-(d_1(n)+kd_k(n))}{\sigma}\left(1+O\left(\frac{k^2}{\sigma}\right)\right)\\ &=\frac{n}{\sigma(k-1)^2}\left(k(k-\varepsilon k-2+\varepsilon)(\alpha+1)-(k(k-2)-\alpha)\right)\left(1+O\left(\frac{k^2}{\sigma}\right)\right)\\ &=\frac{n}{\sigma(k-1)}\left((\alpha-\varepsilon\alpha-\varepsilon)k-\alpha\right)\left(1+O\left(\frac{k^2}{\sigma}\right)\right).\end{aligned}$$

上の式おいて，2 番目の等号は，$0<\varepsilon<\min\left\{\frac{3\alpha}{4(\alpha+1)},\ \frac{1}{12}\right\}$，$4\leq k\leq n^{\frac{1}{8}}$および$0<\alpha<1/2$から得られる．また，これらの条件から最後の式において，$(\alpha-\varepsilon\alpha-\varepsilon)k>\alpha$である．

この結果から，

$$P\left(X_{t_0}\geq \frac{(\alpha-\varepsilon\alpha-\varepsilon)k-\alpha}{k-1}\frac{n}{\sigma}\left(1+O\left(\frac{k^2}{\sigma}\right)\right)\right)\geq 1-\left(\frac{\varepsilon^2 n^{5/16-\delta}}{2}\right)^{-1/2} e^{-\varepsilon^2 n^{5/16-\delta}/6}$$

となるので，ほとんどすべてのコンフィギュレーションF_{t_0}において，$X_{t_0}>h$が成り立つ．

□

次に，X_tの期待値$E(X_t)$と分散$V(X_t)$の大きさを評価する補題を導く．

補題 3

εを$0<\varepsilon<\min\left\{\frac{3\alpha}{4(\alpha+1)},\ \frac{1}{12}\right\}$を満たす任意の定数とする．このとき，以下の式が成り立つ．

$t_0\leq t\leq t_0+\frac{1}{2}h$に対して，

$$\begin{aligned}&M-2t-d_1(n)\left(1-\frac{2t}{M}\right)^{1/2}\left(1+O\left(\frac{1}{n}\right)\right)-kd_k(n)\left(\frac{M-2t}{M-2t_0}\right)^{k/2}\left(1-\frac{(1-\varepsilon)k}{\sigma}\left(1+O\left(\frac{k^2}{\sigma}\right)\right)\right)\\ &\leq E(X_t)\\ &\leq M-2t-d_1(n)\left(1-\frac{2t}{M}\right)^{1/2}\left(1+O\left(\frac{1}{n}\right)\right)-kd_k(n)\left(\frac{M-2t}{M-2t_0}\right)^{k/2}\left(1-\frac{(1+\varepsilon)k}{\sigma}\left(1+O\left(\frac{1}{\sigma}\right)\right)\right),\end{aligned}$$

$$V(X_t)=O\left(\frac{n^2}{\sigma}\right).$$

証明

$X_{t_0}>h$から，t_0におけるオープンコピーの個数はhより大きい．そして，1 ステップごとにオープン

コピーの個数は高々2減少するだけなので，$t_0 \leq t < t_0 + \dfrac{1}{2}h$のとき，$X_t > 0$である．よって，前述の X_t，Y_t および Z_t に関する式(1)は，$I_{\{X_t=0\}}$を含む項を除いて以下のようになる．

$$\begin{aligned}
&M-2t = X_t + Y_t + Z_t,\\
&X_{t+1} = X_t - 2I_{A_{t+1}} - I_{B_{t+1}} + (k-2)I_{C_{t+1}},\\
&Y_{t+1} = Y_t - I_{B_{t+1}},\\
&Z_{t+1} = Z_t - kI_{C_{t+1}}.
\end{aligned} \tag{3}$$

ただし，上記3番目の Y_t に関する式については，$I_{\{X_t=0\}}$を含まないので，$0 \leq t < t_0 + \dfrac{1}{2}h$のときでも成り立つ．また，$t_1 = t_0 + \dfrac{1}{2}h$とすると，$t_0 \leq t \leq t_1 - 1$で上式は有効であり．$Z_t$ について，$Z_{t_0+1} = Z_{t_0} - kI_{C_{t_0+1}}, \cdots, Z_{t_1} = Z_{t_1-1} - kI_{C_{t_1}}$，すなわち，$E(Z_{t_1})$ は有効である．

よって，X_t に対して期待値を，Y_t および Z_t に対して条件付期待値を求めると，以下を得る．

$$\begin{aligned}
&E(X_t) = M - 2t - E(Y_t) - E(Z_t).\\
&E(Y_t | F_{t-1}) = E(Y_{t-1} - I_{B_t} | F_{t-1}),\\
&E(Z_t | F_{t-1}) = E(Z_{t-1} - kI_{C_t} | F_{t-1}).
\end{aligned}$$

後の2つの式のそれぞれ右辺を展開すると，

$$\begin{aligned}
&E(Y_t | F_{t-1}) = Y_{t-1} - P(B_t | F_{t-1}),\\
&E(Z_t | F_{t-1}) = Z_{t-1} - kP(C_t | F_{t-1}).
\end{aligned}$$

したがって，それぞれに対して前述の式(2)を適用して，両辺の期待値をとると，期待値の性質から以下を得る．

$$E(Y_t) = E(Y_{t-1}) - \frac{E(Y_{t-1})}{M-2t+1} = \left(1 - \frac{1}{M-2t+1}\right)E(Y_{t-1}),$$

$$E(Z_t) = E(Z_{t-1}) - k\frac{E(Z_{t-1})}{M-2t+1} = \left(1 - \frac{k}{M-2t+1}\right)E(Z_{t-1}).$$

すなわち，上記2つの式は以下のようになる．

$$\begin{aligned}
&E(Y_t) = \prod_{r=1}^{t}\left(1 - \frac{1}{M-2r+1}\right)E(Y_0) = d_1(n)\prod_{r=1}^{t}\left(1 - \frac{1}{M-2r+1}\right),\\
&E(Z_t) = \prod_{r=t_0+1}^{t}\left(1 - \frac{k}{M-2r+1}\right)E(Z_{t_0}).
\end{aligned}$$

なお，$X_{t_0+(1/2)h-1} > 0$ なので，端点における $E(X_{t_0+(1/2)h})$ は，これらの式から得られることに注意．

ここで，Z_{t_0} は補題2の証明で定義した $R_k^{(t_0)}$を用いて，$Z_{t_0} = kd_k(n) - kR_k^{(t_0)}$と表せるので，ほとんどすべてのコンフィギュレーション F_{t_0} において以下のような不等式が成り立つ．

$$kd_k(n)\left(1 - \frac{(1+\varepsilon)k}{\sigma}\left(1 + O\left(\frac{1}{\sigma}\right)\right)\right) \leq kd_k(n) - kR_k^{(t_0)} \leq kd_k(n)\left(1 - \frac{(1-\varepsilon)k}{\sigma}\left(1 + O\left(\frac{k^2}{\sigma}\right)\right)\right).$$

よって，

$$kd_k(n)\left(1-\frac{(1+\varepsilon)k}{\sigma}\left(1+O\left(\frac{1}{\sigma}\right)\right)\right)\le E(Z_{t_0})\le kd_k(n)\left(1-\frac{(1-\varepsilon)k}{\sigma}\left(1+O\left(\frac{k^2}{\sigma}\right)\right)\right).$$

次に，前述の $\prod_{r=1}^{t}\left(1-\frac{1}{M-2r+1}\right)$ および $\prod_{r=t_0+1}^{t}\left(1-\frac{k}{M-2r+1}\right)$ の大きさを評価すると，以下のようになる．後者の式から，

$$\begin{aligned}
&\log\prod_{r=t_0+1}^{t}\left(1-\frac{k}{M-2r+1}\right)=\sum_{r=t_0+1}^{t}\log\left(1-\frac{k}{M-2r+1}\right)\\
&\le\int_{t_0}^{t}\log\left(1-\frac{k}{M-2x+1}\right)dx=\int_{t_0}^{t}\log(M-2x-k+1)\,dx-\int_{t_0}^{t}\log(M-2x+1)\,dx\\
&=\left[-\frac{(M-2x-k+1)}{2}\log(M-2x-k+1)-x\right]_{t_0}^{t}-\left[-\frac{(M-2x+1)}{2}\log(M-2x+1)-x\right]_{t_0}^{t}\\
&=\log\frac{(1+k/(M-2t-k+1))^{(M-2t-k+1)/2}}{((M-2t_0+1)/(M-2t+1))^{k/2}(1+k/(M-2t_0-k+1))^{(M-2t_0-k+1)/2}}.
\end{aligned}$$

さらに，

$$\begin{aligned}
\log\prod_{r=t_0+1}^{t}\left(1-\frac{k}{M-2r+1}\right)&=\sum_{r=t_0+1}^{t}\log\left(1-\frac{k}{M-2r+1}\right)\\
&\ge\int_{t_0+1}^{t+1}\log\left(1-\frac{k}{M-2x+1}\right)dx\\
&=\int_{t_0+1}^{t+1}\log(M-2x-k+1)\,dx-\int_{t_0+1}^{t+1}\log(M-2x+1)\,dx\\
&=\log\frac{(1+k/(M-2t-k-1))^{(M-2t-k-1)/2}}{((M-2t_0-1)/(M-2t-1))^{k/2}(1+k/(M-2t_0-k-1))^{(M-2t_0-k-1)/2}}.
\end{aligned}$$

よって，以下の不等式を得る．

$$\begin{aligned}
&\frac{(1+k/(M-2t-k-1))^{(M-2t-k-1)/2}}{((M-2t_0-1)/(M-2t-1))^{k/2}(1+k/(M-2t_0-k-1))^{(M-2t_0-k-1)/2}}\\
&\le\prod_{r=t_0+1}^{t}\left(1-\frac{k}{M-2r+1}\right)\\
&\le\frac{(1+k/(M-2t-k+1))^{(M-2t-k+1)/2}}{((M-2t_0+1)/(M-2t+1))^{k/2}(1+k/(M-2t_0-k+1))^{(M-2t_0-k+1)/2}}.
\end{aligned}$$

さらに上式から，

$$\begin{aligned}
&\frac{(1+1/(M-2t-2))^{(M-2t-2)/2}}{((M-1)/(M-2t-1))^{1/2}(1+1/(M-2))^{(M-2)/2}}\\
&\le\prod_{r=1}^{t}\left(1-\frac{1}{M-2r+1}\right)\\
&\le\frac{(1+1/(M-2t))^{(M-2t)/2}}{((M+1)/(M-2t+1))^{1/2}(1+1/M)^{M/2}}.
\end{aligned}$$

ここで，上記の $\dfrac{(1+k/(M-2t-k+1))^{(M-2t-k+1)/2}}{(1+k/(M-2t_0-k+1))^{(M-2t_0-k+1)/2}}$ の大きさを以下のように評価する．

$m=M-2t-k+1$ および $a=2t-2t_0$ とすると，$\dfrac{n}{2}<m<\dfrac{3}{2}n$，$0\leq a<n$ である．このとき，以下となる．

$$r=\frac{(1+k/(M-2t-k+1))^{(M-2t-k+1)/2}}{(1+k/(M-2t_0-k+1))^{(M-2t_0-k+1)/2}}=\frac{(1+k/m)^{m/2}}{(1+k/(m+a))^{(m+a)/2}}\leq 1.$$

$$\begin{aligned}\log r&=\frac{m}{2}\log\Big(1+\frac{k}{m}\Big)-\frac{m+a}{2}\log\Big(1+\frac{k}{m+a}\Big)\\&=\frac{m}{2}\sum_{i=1}^{\infty}\frac{(-1)^{i+1}}{i}\Big(\frac{k}{m}\Big)^i-\frac{m+a}{2}\sum_{i=1}^{\infty}\frac{(-1)^{i+1}}{i}\Big(\frac{k}{m+a}\Big)^i\\&\geq\frac{m}{2}\sum_{i=2}^{\infty}\frac{(-1)^{i+1}}{i}\Big(\frac{k}{m}\Big)^i\geq-\frac{m}{4}\Big(\frac{k}{m}\Big)^2=-\frac{k^2}{4m}.\end{aligned}$$

よって，$r=1+O\Big(\dfrac{k^2}{n}\Big)$．上記不等式における他の式も同様に考えられるので，結局，以下を得る．

$$\prod_{r=1}^{t}\Big(1-\frac{1}{M-2r+1}\Big)=\Big(1-\frac{2t}{M}\Big)^{1/2}\Big(1+O\Big(\frac{1}{n}\Big)\Big).$$

$$\prod_{r=t_0+1}^{t}\Big(1-\frac{k}{M-2r+1}\Big)=\Big(\frac{M-2t}{M-2t_0}\Big)^{k/2}\Big(1+O\Big(\frac{k^2}{\sigma}\Big)\Big).$$

したがって，上述の $E(Y_t)$ および $E(Z_t)$ は，以下のように評価できる．

$$E(Y_t)=d_1(n)\Big(1-\frac{2t}{M}\Big)^{1/2}\Big(1+O\Big(\frac{1}{n}\Big)\Big).$$

$$\begin{aligned}&kd_k(n)\Big(\frac{M-2t}{M-2t_0}\Big)^{k/2}\Big(1-\frac{(1+\varepsilon)k}{\sigma}\Big(1+O\Big(\frac{1}{\sigma}\Big)\Big)\Big)\leq E(Z_t)\\&\qquad\leq kd_k(n)\Big(\frac{M-2t}{M-2t_0}\Big)^{k/2}\Big(1-\frac{(1-\varepsilon)k}{\sigma}\Big(1+O\Big(\frac{k^2}{\sigma}\Big)\Big)\Big).\end{aligned}$$

よって，$E(X_t)=M-2t-E(Y_t)-E(Z_t)$ なので，以下のように主張の式を得る．

$$\begin{aligned}&M-2t-d_1(n)\Big(1-\frac{2t}{M}\Big)^{1/2}\Big(1+O\Big(\frac{1}{n}\Big)\Big)-kd_k(n)\Big(\frac{M-2t}{M-2t_0}\Big)^{k/2}\Big(1-\frac{(1-\varepsilon)k}{\sigma}\Big(1+O\Big(\frac{k^2}{\sigma}\Big)\Big)\Big)\\&\leq E(X_t)\\&\leq M-2t-d_1(n)\Big(1-\frac{2t}{M}\Big)^{1/2}\Big(1+O\Big(\frac{1}{n}\Big)\Big)-kd_k(n)\Big(\frac{M-2t}{M-2t_0}\Big)^{k/2}\Big(1-\frac{(1+\varepsilon)k}{\sigma}\Big(1+O\Big(\frac{1}{\sigma}\Big)\Big)\Big).\end{aligned}$$

次に，$V(X_t)$ の大きさを評価する．まず，式(3)から以下の式が成り立つ．

$$V(X_t)=E(X_t^2)-(E(X_t))^2=E((M-2t-Y_t-Z_t)^2)-(M-2t-E(Y_t)-E(Z_t))^2$$

$$=E(Y_t^2)-(E(Y_t))^2+E(Z_t^2)-(E(Z_t))^2+2(E(Y_tZ_t)-E(Y_t)E(Z_t)).$$

ここで，式(2)および式(3)から，以下を得る．

$$E(Y_t^2|F_{t-1})=E(Y_{t-1}{}^2-2Y_{t-1}I_{B_t}+I_{B_t}|F_{t-1})$$
$$=Y_{t-1}{}^2-2Y_{t-1}P(B_t|F_{t-1})+P(B_t|F_{t-1})=Y_{t-1}{}^2\left(1-\frac{2}{M-2t+1}\right)+\frac{Y_{t-1}}{M-2t+1}.$$

この式の両辺の期待値をとって，

$$E(Y_t^2)=\left(1-\frac{2}{M-2t+1}\right)E(Y_{t-1}{}^2)+\frac{E(Y_{t-1})}{M-2t+1}.$$

ここで，前に得られた式

$$E(Y_t)=\left(1-\frac{1}{M-2t+1}\right)E(Y_{t-1})$$

を用いて各辺同士を引くと，以下を得る．

$$E(Y_t^2)-E(Y_t)=\left(1-\frac{2}{M-2t+1}\right)(E(Y_{t-1}{}^2)-E(Y_{t-1})).$$

よって，

$$E(Y_t^2)=\prod_{r=1}^{t}\left(1-\frac{2}{M-2r+1}\right)(E(Y_0{}^2)-E(Y_0))+E(Y_t)=(d_1(n))^2\left(1-\frac{2t}{M}\right)\left(1+O\left(\frac{1}{n}\right)\right).$$

したがって，

$$E(Y_t^2)-(E(Y_t))^2=(d_1(n))^2\left(1-\frac{2t}{M}\right)O\left(\frac{1}{n}\right).$$

同様に Z_t について考えるが，Z_t に関する式(1)は，$t_0\leq t< t_0+\frac{1}{2}h$ において $X_t>0$ なので，このとき，$I_{\{X_t=0\}}$ を含む項を除けることに注意する．したがって，以下を得る．

$$E(Z_t^2)-kE(Z_t)=\prod_{r=t_0+1}^{t}\left(1-\frac{2k}{M-2r+1}\right)(E(Z_{t_0}{}^2)-kE(Z_{t_0})).$$

よって，

$$E(Z_t^2)=\left(1-\frac{2t}{M}\right)^k\left(1+O\left(\frac{k}{\sigma}\right)\right)(E(Z_{t_0}{}^2)-kE(Z_{t_0}))+kE(Z_t)$$
$$=\left(1-\frac{2t}{M}\right)^k\left(1+O\left(\frac{k}{\sigma}\right)\right)\left((kd_k(n))^2\left(1+O\left(\frac{k}{\sigma}\right)\right)-k^2d_k(n)\left(1+O\left(\frac{k}{\sigma}\right)\right)\right)$$
$$+k^2d_k(n)\left(1-\frac{2t}{M}\right)^{\frac{k}{2}}\left(1+O\left(\frac{k^3}{\sigma^2}\right)\right)$$
$$=(kd_k(n))^2\left(1-\frac{2t}{M}\right)^k\left(1+O\left(\frac{k}{\sigma}\right)\right).$$

これから，

$$E(Z_t^2)-(E(Z_t))^2=(kd_k(n))^2\left(1-\frac{2t}{M}\right)^k O\left(\frac{k}{\sigma}\right).$$

次に，$Y_t Z_t$ について，$Y_t Z_t=(Y_{t-1}-I_{B_t})(Z_{t-1}-kI_{C_t})$ なので，

$$E(Y_t Z_t|F_{t-1})=Y_{t-1}Z_{t-1}-kY_{t-1}P(C_t|F_{t-1})-Z_{t-1}P(B_t|F_{t-1})+kP(B_t\cap C_t|F_{t-1})$$

となり，この式の両辺の期待値をとることにより以下を得る.

$$E(Y_t Z_t)=\prod_{r=t_0+1}^{t}\left(1-\frac{k+1}{M-2r+1}\right)E(Y_{t_0}Z_{t_0}).$$

したがって，

$$E(Y_t Z_t)-E(Y_t)E(Z_t)=kd_1(n)d_k(n)\left(1-\frac{2t}{M}\right)^{\frac{k+1}{2}}O\left(\frac{k}{\sigma}\right)$$

となるので，以下を得る.

$$V(X_t)=O\left(\frac{n^2}{\sigma}\right).$$

□

この補題から，X_t は期待値付近に分布すると予想できる．そこで，実際に X_t はその期待値付近に分布することを示す．この証明に当たっては，Azuma—Hoeffding bound（例えば，Frize and Karonski（2016）を参照）を適用する.

補題 4

コンフィグレーションが第 t_0 ステップまで終了しているとする．そして，t_0 からコンフィグレーションを続行するとして，$t_0\leq t<t_1$ を満たす任意の t に対し，$X_t>0$ とする．また，a は任意の正の定数とする．このとき，以下の式が成り立つ.

$$t_0\leq t\leq t_1 \text{ に対して，} P(|X_t-E(X_t)|\geq a)\leq 2\exp\left(-\frac{a^2}{8(t-t_0)k^2}\right).$$

証明

いま，最初の時点を t_0 とし，$X_{t_0}=N$ とする．そして，$t_0\leq t<t_1=N/2$ とすれば，$X_t>0$ であり，$E(X_t)$ の大きさは補題 3 のように評価できる．次に，事象 H_i を第 i 回目のステップにおけるコピー（後で選択される方）のランダムな選択とする．そして，一連の事象を $H=H_{t_0+1},\cdots,H_t$ とし，ある関数 f に対して，$X_t=f(H)$ とする．このとき，$t_0+1\leq i\leq t$ に対して，以下の式の上限を評価する.

$$|E(X_t|H_{t_0+1},\cdots,H_i)-E(X_t|H_{t_0+1},\cdots,H_{i-1})|.$$

そこで，i 番目のステップまで実行したとする．すなわち，各ステップにおいて最初に 1 個のコピーを任意に選択し，次にもう 1 個のコピーをランダムに選択するので，$2i$ 個のコピーまで選択されたことになる．L はこの時点における未選択なコピーの集合とする．L の大きさは，$M-2i$ である．そして，事象 $H_{t_0+1},\cdots,H_{i-1}$ においてそれぞれコピー$a_{t_0+1},\cdots,a_{i-1}$ が選択され，かつ事象 H_i においてコピーx が選択されるとき，X_t の条件付期待値 $E_i(x)$ を以下のように定義する.

$$E_i(x) = E(X_t \mid H_{t_0+1}, \cdots, H_i).$$

また，$H_{t_0+1}, \cdots, H_{i-1}$ においてそれぞれコピー$a_{t_0+1}, \cdots, a_{i-1}$ が選択される条件下で，H_i においてコピーx の選択される条件付確率を $P_{H_i}(x)$ と記す．このとき，$H_{t_0+1}, \cdots, H_{i-1}$ においてそれぞれコピー$a_{t_0+1}, \cdots, a_{i-1}$ が選択され，かつ H_i においてコピーb が選択される条件下で，X_t の期待値 $E_i(b)$ は明らかに以下の式を満たす．

$$E_i(b) = E(X_t \mid H_{t_0+1}, \cdots, H_i) = \mathrm{P}_{H_i}(b) \times E_i(b) + \Big(\sum_{c \neq b} \mathrm{P}_{H_i}(c)\Big) \times E_i(b).$$

ただし，c は H_i においてランダムに選択される $c \neq b$ となるコピーである．

一方，期待値の性質から以下の式を得る．

$$E(X_t \mid H_{t_0+1}, \cdots, H_{i-1}) = \mathrm{P}_{H_i}(b) \times E_i(b) + \sum_{c \neq b} \mathrm{P}_{H_i}(c) \times E_i(c).$$

よって，上の 2 つの式から以下の不等式を得る．

$$|E(X_t \mid H_{t_0+1}, \cdots, H_i) - E(X_t \mid H_{t_0+1}, \cdots, H_{i-1})| \leq \sum_{c \neq b} \mathrm{P}_{H_i}(c) \times |E_i(b) - E_i(c)|.$$

このとき，前述の $\mu(x)$ を用いれば，

$$E_i(b) = \mu(a_{t_0+1}) + \cdots + \mu(a_{i-1}) + \mu(b) + E\Big(\sum_{j=i+1}^{t} Q_j\Big),$$

$$E_i(c) = \mu(a_{t_0+1}) + \cdots + \mu(a_{i-1}) + \mu(c) + E\Big(\sum_{j=i+1} Q'_j\Big)$$

である．ただし，確率変数 Q_j (Q'_j) は，$H_{t_0+1}, \cdots, H_{i-1}$ においてそれぞれ $a_{t_0+1}, \cdots, a_{i-1}$ が選択され（Q'_j の場合も同じ)，ステップ i において b が選択される（Q'_j の場合は c が選択）という条件下で，j 番目のステップにおいて選択されるコピーx の $\mu(x)$ 値をとる．このことから，以下を得る．

$$|E_i(b) - E_i(c)| \leq |\mu(b) - \mu(c)| + \Big|E\Big(\sum_{j=i+1}^{t} Q_j\Big) - E\Big(\sum_{j=i+1}^{t} Q'_j\Big)\Big|.$$

そこで，i において b を選択して生成される途中段階のコンフィギュレーションを F_i とし，i において c を選択して生成されるそれを F'_i とする（$i-1$ 以前のコピー対はともに同じ)．そして，b のもとになる頂点を u とし，c のもとになる頂点を v とするとき，u と v のコピー全てからなる集合を K とする．ここで，F_i と F'_i に対して，それぞれ $i+1$ から t までステップを続行するとき，前者から K を除いたコピー全体および後者から K を除いたコピー全体では，i 以前にできたコピー対はそれぞれ同じで，$i+1$ から t まで各コピーx はランダム一様に選択されるから $\mu(x)$ の和の期待値はそれぞれ同じである．一方，F_i が構成されるプロセスにおける K と F'_i が構成されるプロセスにおける K について同様に考えると（両者の場合，$X_{t_0} = N$ かつ $t_0 \leq t < t_1 = N/2$ なので，$t_0 \leq t \leq t_1$ では，前述に示したコピーの任意選択の定義から任意選択は行われないことに注意)，F_i における K では，b は既に選択されており，c を含む未選択のコピーがランダム一様に選択され，F'_i における K では，c が既に選択されており，b を含む未選択のコピーがランダム一様に選択されるので，$\mu(x)$ の和の期待値の差は高々k となる．し

たがって，以下を得る．

$$|E_i(b)-E_i(c)|\leq 2k.$$

すなわち，

$$|E(X_t|H_{t_0+1},\cdots,H_i)-E(X_t|H_{t_0+1},\cdots,H_{i-1})|\leq 2k.$$

このとき，$Z_i=E(X_t|H_{t_0+1},\cdots,H_i)$とすると，$|Z_i-Z_{i-1}|\leq 2k\quad(t_0+1\leq i\leq t)$ であり，$\{Z_i\}_{t_0}^{t}$ は H_{t_0+1}, $\cdots, H_t$ に関して Martingale なので，Azuma—Hoeffding bound によって，任意の $a>0$ に対して，以下の不等式が成り立つ．

$$t_0\leq t\leq t_1 \text{ に対して，}\quad P(|Z_t-E(Z_t)|\geq a)\leq 2\exp\left(-\frac{a^2}{8(t-t_0)k^2}\right).$$

ここで，$Z_t=X_t$ であるから主張を得る． □

以上述べてきたこれらの補題から，コンフィグレーションによって構成されるランダムグラフの大きさに関する以下の定理 1 を導出できる．

定理 1

ある定数 $\zeta>0$ に対して，ほとんどすべてのランダムな単純グラフ G は，少なくとも$\zeta\dfrac{n}{k^2}$個の頂点をもつ連結成分を含む．

証明

$1\leq i\leq m$ に対して，$t_i=\left(1+\dfrac{\alpha n}{4M}\right)t_{i-1}$ とすると，$t_m=\left(1+\dfrac{\alpha n}{4M}\right)^m t_0$ である．そして，$t_m\leq\dfrac{\alpha M}{10k}$ とする．このとき，ほとんどすべての F_t において，以下が成り立つことを帰納法によって示す．

$1\leq t\leq t_0$ においてバックエッジの起こらない事象を $\bar{B}$ とし，また，$0\leq i\leq m$ とし，$t_0\leq t<t_i$ に対して $X_t>0$ であり，かつ，$X_{t_i}>\dfrac{\alpha n}{2M}t_i$ となる事象を E_i とすると，定数 $0<\mu<1/6$ に対して，

$$P(E_0\cap\cdots\cap E_i|\bar{B})\geq 1-2(i+1)\exp\left(-\frac{1}{200}\left(\frac{9\alpha\mu}{4+\alpha}\right)^2 n^{1/16}\right).$$

まず，$i=0$ の場合，補題 2 から，

$$h=\frac{9((\alpha-\varepsilon\alpha-\varepsilon)k-\alpha)}{10(k-1)}\frac{n}{\sigma}>\frac{\alpha}{2}\frac{n}{\sigma}=\frac{\alpha n}{2M}t_0\quad \text{かつ}$$

$$P(X_{t_0}>h|\bar{B})\geq 1-\left(\frac{\varepsilon^2 n^{5/16-\delta}}{2}\right)^{-1/2}e^{-\varepsilon^2 n^{5/16-\delta}/6}$$

となるから，主張は成り立つ．

次に，$i-1$ の場合，主張が成り立つと仮定する．そして，

$$P(E_0\cap\cdots\cap E_i|\bar{B})=P(E_0\cap\cdots\cap E_{i-1}|\bar{B})P(E_i|E_0\cap\cdots\cap E_{i-1}\cap\bar{B})$$

となるので，以下では，$P(E_i|E_0\cap\cdots\cap E_{i-1}\cap\bar{B})$すなわち，$E_0\cap\cdots\cap E_{i-1}\cap\bar{B}$ が成り立つという条件のもとで，E_i の起こる確率の下限を評価する．

X_t は 1 ステップごとに高々2 減少するにすぎないから，前提条件の E_{i-1} によって，第 t_0 ステップから第 $(t_{i-1}+(1/2)X_{t_{i-1}}-1)$ステップまで $X_t>0$ である．すなわち，$t_0\leq t\leq t_{i-1}+(\alpha n/4M)t_{i-1}-1=$

t_i-1 において，$X_t>0$ となる．これは，E_i の前半部分である．このとき，式(3)が成り立つので，補題 3 と同様に考えて以下を得る．なお，$0<\varepsilon<\dfrac{\alpha}{80}<\min\left\{\dfrac{3\alpha}{4(\alpha+1)}, \dfrac{1}{12}\right\}$に注意.

$$
\begin{aligned}
E(X_t) \geq\ & M-2t-d_1(n)\left(1-\frac{2t}{M}\right)^{1/2}\left(1+O\left(\frac{1}{n}\right)\right) \\
& -kd_k(n)\left(\frac{M-2t}{M-2t_0}\right)^{k/2}\left(1-\frac{(1-\varepsilon)k}{\sigma}\left(1+O\left(\frac{k^2}{\sigma}\right)\right)\right) \\
\geq\ & M-2t-d_1(n)e^{-t/M}\left(1+O\left(\frac{1}{n}\right)\right) \\
& -kd_k(n)\left(1-\frac{2t}{M}\right)^{k/2}\frac{1}{(1-2t_0/M)^{k/2}}\left(1-\frac{(1-\varepsilon)k}{\sigma}\left(1+O\left(\frac{k^2}{\sigma}\right)\right)\right) \\
\geq\ & M-2t-d_1(n)\left(1-\frac{t}{M}+\frac{1}{2}\left(\frac{t}{M}\right)^2\right)\left(1+O\left(\frac{1}{n}\right)\right) \\
& -kd_k(n)e^{-kt/M}\frac{1}{1-kt_0/M}\left(1-\frac{(1-\varepsilon)k}{\sigma}\left(1+O\left(\frac{k^2}{\sigma}\right)\right)\right) \\
\geq\ & M-2t-d_1(n)\left(1-\frac{t}{M}+\frac{1}{2}\left(\frac{t}{M}\right)^2\right)\left(1+O\left(\frac{1}{n}\right)\right) \\
& -kd_k(n)\left(1-\frac{kt}{M}+\frac{1}{2}\left(\frac{kt}{M}\right)^2\right)\left(1+(1+\varepsilon)\frac{k}{\sigma}\right)\left(1-\frac{(1-\varepsilon)k}{\sigma}\left(1+O\left(\frac{k^2}{\sigma}\right)\right)\right) \\
\geq\ & M-2t-d_1(n)\left(1-\frac{t}{M}+\left(\frac{t}{M}\right)^2\right)-kd_k(n)\left(1-\frac{kt}{M}+\frac{1}{2}\left(\frac{kt}{M}\right)^2\right)\left(1+2\varepsilon\frac{k}{\sigma}+O\left(\frac{k^3}{\sigma^2}\right)\right) \\
\geq\ & M-2t-d_1(n)\left(1-\frac{t}{M}+\left(\frac{t}{M}\right)^2\right)-kd_k(n)\left(1-\frac{kt}{M}+\frac{1}{2}\left(\frac{kt}{M}\right)^2\right)\left(1+3\varepsilon\frac{k}{\sigma}\right) \\
\geq\ & M-2t-d_1(n)\left(1-\frac{t}{M}\right)-kd_k(n)\left(1-\frac{kt}{M}\right) \\
& -\left(d_1(n)\left(\frac{t}{M}\right)^2+kd_k(n)\left(3\varepsilon\frac{k}{\sigma}+\frac{1}{2}\left(\frac{kt}{M}\right)^2+\frac{3\varepsilon}{2}\left(\frac{kt}{M}\right)^2\frac{k}{\sigma}\right)\right) \\
\geq\ & \frac{\alpha nt}{M}-\left(d_1(n)\left(\frac{t}{M}\right)^2+kd_k(n)\left(3\varepsilon\frac{k}{\sigma}+\left(\frac{kt}{M}\right)^2\right)\right)\geq\frac{3\alpha nt}{5M}.
\end{aligned}
$$

ここで，最初から 4 番目の不等式では，$t\leq\dfrac{\alpha M}{10k}$ のとき，$e^{-kt/M}\leq 1-\dfrac{kt}{M}+\dfrac{1}{2}\left(\dfrac{kt}{M}\right)^2$ および $\dfrac{1}{1-kt_0/M}\leq 1+(1+\varepsilon)\dfrac{k}{\sigma}$ が成り立つことを用い，最後から 2 番目の不等式では，$M=d_1(n)+kd_k(n)$ および $k(k-2)d_k(n)-d_1(n)=\alpha n$ を用いている．また，最後の不等式は以下のように評価できることによる．

$$d_1(n)\left(\frac{t}{M}\right)^2=\frac{td_1(n)}{M}\frac{t}{M}\leq\frac{\alpha d_1(n)}{10k}\frac{t}{M}\leq\frac{\alpha nt}{40M},$$

$$3\varepsilon\frac{k^2d_k(n)}{\sigma}\leq\frac{M}{t_0}\frac{3\varepsilon k^2d_k(n)}{\sigma}\frac{t}{M}\leq\sigma\frac{3\varepsilon}{\sigma}\frac{8n}{3}\frac{t}{M}\leq\frac{\alpha nt}{10M}\text{および},$$

$$\left(\frac{kt}{M}\right)^2kd_k(n)=\left(\frac{kt}{M}k^2d_k(n)\right)\frac{t}{M}\leq\frac{8\alpha nt}{30M}.$$

なお，$k^2d_k(n)=(\alpha+1)\left(\frac{k}{k-1}\right)^2n\leq\frac{8n}{3}$であることに注意．上記の式は，$t=t_i$ の場合も成り立つので，$E(X_{t_i})\geq\frac{3\alpha nt_i}{5M}$である．

ところで，補題 4 において，$t=t_i$ および $t_0=t_{i-1}$，そして $0<\mu<1/6$ とし，$a=\mu E(X_{t_i})$とすると，この結果を用いて，

$$\begin{aligned}P(|X_{t_i}-E(X_{t_i})|\leq\mu E(X_{t_i}))&\geq1-2\exp\left(-\frac{(\mu E(X_{t_i}))^2}{8(t_i-t_{i-1})k^2}\right)\\&\geq1-2\exp\left(-\frac{\mu^2\left(\frac{3\alpha nt_i}{5M}\right)^2}{8(t_i-t_{i-1})k^2}\right)\\&\geq1-2\exp\left(-\frac{\mu^2}{8}\left(\frac{9\alpha}{5(4+\alpha)}\right)^2\frac{t_i}{k^2}\right)\\&\geq1-2\exp\left(-\frac{1}{200}\left(\frac{9\alpha\mu}{4+\alpha}\right)^2\frac{t_0}{k^2}\right)\\&\geq1-2\exp\left(-\frac{1}{200}\left(\frac{9\alpha\mu}{4+\alpha}\right)^2n^{1/16}\right).\end{aligned}$$

なお，上式の 3 番目の不等式は，$k\geq4$ かつ $M=\frac{k+\alpha}{k-1}n$ による。よって，少なくとも確率 $1-2\exp\left(-\frac{1}{200}\left(\frac{9\alpha\mu}{4+\alpha}\right)^2n^{1/16}\right)$において，以下の式が成り立つ．

$$X_{t_i}\geq(1-\mu)E(X_{t_i})\geq(1-\mu)\frac{3\alpha nt_i}{5M}>\frac{\alpha nt_i}{2M}.$$

上述のことから，以下を得る．

$$P(E_i|E_0\cap\cdots\cap E_{i-1}\cap\overline{B})\geq1-2\exp\left(-\frac{1}{200}\left(\frac{9\alpha\mu}{4+\alpha}\right)^2n^{1/16}\right).$$

したがって，

$$\begin{aligned}P(E_0\cap\cdots\cap E_i|\overline{B})&\geq\left(1-2i\exp\left(-\frac{1}{200}\left(\frac{9\alpha\mu}{4+\alpha}\right)^2n^{1/16}\right)\right)\left(1-2\exp\left(-\frac{1}{200}\left(\frac{9\alpha\mu}{4+\alpha}\right)^2n^{1/16}\right)\right)\\&\geq1-2(i+1)\exp\left(-\frac{1}{200}\left(\frac{9\alpha\mu}{4+\alpha}\right)^2n^{1/16}\right).\end{aligned}$$

一方，補題 1 から，

$$P(\overline{B}) \geq 1 + O\left(\frac{1}{n^{2\delta}}\right).$$

したがって，

$$P(E_0 \cap \cdots \cap E_m \cap \overline{B}) = P(\overline{B})P(E_0 \cap \cdots \cap E_m | \overline{B}) \geq 1 + O\left(\frac{m}{n^{2\delta}}\right).$$

ここで，$t_m = \left(1 + \frac{\alpha}{4}\frac{n}{M}\right)^m t_0 \leq \frac{\alpha M}{10k}$ なので，$t_0 = \frac{M}{\sigma}$ および $\sigma = n^{\frac{9}{16}+\delta}\left(0 < \delta < \frac{1}{8}\right)$ を用いれば，m の上限を評価する以下の式を得る．

$$m \leq \frac{\log(\alpha\sigma/(10k))}{\log(1 + \alpha n/(4M))} \leq \frac{(\delta + 9/16)\log n}{\log(1 + 3\alpha/(4(\alpha + 4)))}.$$

このことから，ほとんどすべての F_t において，E_m が成り立つ．そして，m はこの上限までとれるから，第 t_m ステップにおいて，少なくとも $\frac{\alpha n}{2M}t_m = \frac{\alpha^2 n}{20k}$ 個のオープンコピーがある．k 個のオープンコピーにつき少なくとも 1 個の頂点が対応するから，ある定数 $\zeta > 0$ に対して，ほとんどすべての G は，少なくとも $\zeta\frac{n}{k^2}$ 個の頂点をもつ連結成分を含むことが示される． □

前述の議論では，補題 1 を満たす σ に対して，$t_0 = M/\sigma$ としたが，t_0 をもっと小さくしても，近似の部分と事象の成り立つ確率に注意すれば，定理 1 に至る前述の一連の議論は成り立つ．よって，以下の定理 2 を提示する．

定理 2

λ は正の任意の定数とする．このとき，$k^2(\log n)^\lambda \leq t \leq \frac{\alpha M}{10k}$ では，ほとんどすべてのコンフィギュレーション F_t において，オープンコピーの数は 0 にならない．

証明

$t_0 = k^2(\log n)^\lambda$ すなわち，$\sigma = M/(k^2(\log n)^\lambda)$ とすると，補題 1 から定理 1 に至るそれぞれの証明は成り立つ． □

定理 1 からわかるように，k が定数の場合には，コンフィグレーションによって構成される連結グラフは，頂点数が n のオーダーである．そうでない場合，最大の頂点数の連結グラフは n/k^2 のオーダーになる可能性がある．これを考察する．

定理 1 の証明からわかるように，X_t は $E(X_t)$ の付近に分布する．よって X_t の分布は，ほぼ次の関数の正の部分によって表されると考えてよい．

$$f(x) = M - 2x - d_1(n)\left(1 - \frac{2x}{M}\right)^{1/2} - kd_k(n)\left(1 - \frac{2x}{M}\right)^{k/2}.$$

そこで，この関数の特徴を以下の補題として示す．

補題 5

$n\to\infty$ のとき，$k\to\infty$ とする．以下の関数 $f(x)$ の最大値を与える点の x 座標は，k の増加とともに原点側に動く．$0\leq x\leq M/2$ に対して，

$$f(x)=M-2x-d_1(n)\left(1-\frac{2x}{M}\right)^{1/2}-kd_k(n)\left(1-\frac{2x}{M}\right)^{k/2}.$$

証明

与えられた式から，$f(0)=0$ および $f\left(\frac{M}{2}\right)=0$ である．$f(x)$ を微分すると，

$$f'(x)=-2+\frac{d_1(n)}{M}\left(1-\frac{2x}{M}\right)^{-1/2}+\frac{k^2d_k(n)}{M}\left(1-\frac{2x}{M}\right)^{k/2-1}.$$

よって，$f'(0)=-2+\frac{d_1(n)}{M}+\frac{k^2d_k(n)}{M}>0$ かつ $\lim_{x\uparrow M/2}f'(x)=+\infty$ である．このとき，$g(x)=\frac{d_1(n)}{M}\left(1-\frac{2x}{M}\right)^{-1/2}+\frac{k^2d_k(n)}{M}\left(1-\frac{2x}{M}\right)^{k/2-1}$ とすると，$f'(x)=g(x)-2$ なので，$g(0)>2$ かつ $\lim_{x\uparrow M/2}g(x)=+\infty$ である．このとき，$0\leq x\leq M/2$ に対して，$g(x)\geq 2$ とすると，$f'(x)\geq 0$ となり，$f(0)=0$ および $f\left(\frac{M}{2}\right)=0$ という事実に反するので，$g(x)$ は $0\leq x\leq M/2$ のある区間で 2 より小となる．$g(x)$ を微分すると，

$$g'(x)=\frac{d_1(n)}{M^2}\left(1-\frac{2x}{M}\right)^{-3/2}-\frac{k^2(k-2)d_k(n)}{M^2}\left(1-\frac{2x}{M}\right)^{k/2-2}.$$

よって $g'(x)=0$ を満たす x は，以下のように唯 1 つ存在する．これを x_0 とする．

$$0<x_0=\frac{M}{2}\left(1-\left(\frac{1}{k-2}\,\frac{k(k-2)-\alpha}{k^2(\alpha+1)}\right)^{2/(k-1)}\right)<\frac{M}{2}.$$

また，$g'(0)=-\alpha n/M^2<0$ かつ $\lim_{x\uparrow M/2}g'(x)=+\infty$ を考慮すれば，$g(x)$ は $g(x_1)=2$ および $g(x_2)=2$ であるような区間 $x_1<x<x_2$ で 2 より小になる．そして，$x_1<x_0<x_2$ であり，$f(x)$ では x_1 は極大値（最大値），x_0 は変曲点，x_2 は極小値を与える．このとき，x_0 の大きさを以下のように評価できる．

$$x_0=\frac{M}{2}\left(1-\left(\frac{1}{k-2}\,\frac{k(k-2)-\alpha}{k^2(\alpha+1)}\right)^{2/(k-1)}\right)<\frac{M}{2}\left(1-\left(\frac{1}{2(k+1)}\right)^{2/(k-1)}\right).$$

右辺の式は，k を大きくするにつれ 0 に近づくので，x_1 も 0 に近づく． □

k を大きくするにつれ，変曲点は左側に動くので，$y=f(x)$ のグラフも左側に動くと推測できる．すなわち，X_t の最大値，さらには，0 になる点が左側に動くことになる．よって，補題 5 と定理 1 に基づき以下の予想を提示する．

予想

$n\to\infty$ のとき，$k\to\infty$ となる場合，ほとんどすべての G において，その連結成分の頂点数は，ある定数 $\zeta>0$ に対して，せいぜい $\zeta\frac{n}{k^2}$ 個である．

さて補題1にあるように，この区間では，ほとんどすべてのランダムグラフにおいて，2個のオープンコピーのマッチングは行われない．このとき，当区間における頂点コピーの選択は次のような試行と見なせる．なお，一般性のため，初期値は a とする．

$$X_0 = a \quad (a \text{ は正の整数}),$$

$$T_r = \begin{cases} k-2 & : r \text{ 回目に } L \text{ から次数 } k \text{ の未開示な頂点のコピーを選択する場合,} \\ -1 & : r \text{ 回目に } L \text{ から次数 } 1 \text{ の未開示な頂点のコピーを選択する場合,} \end{cases}$$

$$X_r = X_0 + T_1 + \cdots + T_r \quad (r \geq 1).$$

そこで，t 回までの試行において，$X_1 \geq 1$，…，$X_{t-1} \geq 1$，$X_t = 0$ となるような試行の組の総数を近似評価する．まず，最初の状態 a から出発して，途中 X_r $(1 \leq r \leq t-1)$ が 0 以下になる場合も含めて，$X_t = 0$ となるすべての場合の数を $N_t(a,0)$ と記す（試行は t 回まで継続して行う）．次に，j $(1 \leq j \leq t)$ 回目の試行において初めて $X_j = 0$ となるすべての場合の数を $N'_j(a,0)$ と記す．

さて，X_j が 0 になるならば，i を次数 k の未開示な頂点のコピーを選択する回数として，$j = (k-1)i + a$ と表され，逆に，$j = (k-1)i + a$ と表されるならば，$X_j = 0$ となる．よって，$t = (k-1)m + a$ とすると，以下のような漸化式が成り立つ．

$$N_t(a,0) = \sum_{i=0}^{m} N'_{(k-1)i+a}(a,0) \binom{(k-1)(m-i)}{m-i}.$$

この式から帰納的に $N'_t(a,0)$ の大きさを評価することも考えられるが，より簡潔な評価が得られる次の補題のような不等式を導く．

補題 6

a は正の整数とする．負でない整数 i に対して，以下の不等式が成り立つ．

$$N'_a(a,0) = 1,$$

$$a(k-1)^{i-1} \leq N'_{(k-1)i+a}(a,0)$$
$$\leq \binom{(k-1)i+a}{i} - \binom{(k-1)i}{i} - a \sum_{j=1}^{i-1} (k-1)^{j-1} \binom{(k-1)(i-j)}{i-j} \quad (i \geq 1).$$

証明

まず，$N'_{(k-1)(i+1)+a}(a,0)$ と $N'_{(k-1)i+a}(a,0)$ の関係を考える．$(k-1)i+a$ 回目の試行で初めてオープンコピーの数が 0 になる場合は，$(k-1)(i-1)+a$ 回目の試行でオープンコピーの数が $k-1$ 個になる状態 P を常に経由し，その後，ただ1通りの一連の試行を経てオープンコピーの数が 0 になることを意味する（P に至る途中でオープンコピーの数は 0 にならないものとする）．一方，$(k-1)(i+1)+a$ 回目の試行で初めてオープンコピーの数が 0 になる場合は．P を経由する $k-1$ 通りの場合と P を経由しない場合に分けられる．したがって，以下の式を得る．

$$N'_a(a,0) = 1, \quad N'_{k-1+a}(a,0) = a, \quad \text{かつ } i \geq 1 \text{ に対して,}$$

$$(k-1) N'_{(k-1)i+a}(a,0) \leq N'_{(k-1)(i+1)+a}(a,0).$$

よって，主張にある不等式の左側の部分を得る．

次に，X_j が 0 になるのは，ある負でない整数 i を用いて，$j = (k-1)i + a$ と表される場合に限られる．

このとき，L から次数 k の未開示な頂点のコピーを選択する回数は i であり，次数 1 の未開示な頂点のコピーを選択する回数は $(k-2)i+a$ である．よって，$N_{(k-1)i+a}(a,0)$ は以下のように表せる．

$$N_{(k-1)i+a}(a,0)=\binom{(k-1)i+a}{i}.$$

よって，前述の議論における漸化式および先ほど示した主張にある不等式の左側の部分を用いて以下を得る．$i\geq 1$ の場合，

$$N'_{(k-1)i+a}(a,0)=\binom{(k-1)i+a}{i}-\sum_{j=0}^{i-1}N'_{(k-1)j+a}(a,0)\binom{(k-1)(i-j)}{i-j}$$
$$\leq\binom{(k-1)i+a}{i}-\binom{(k-1)i}{i}-a\sum_{j=1}^{i-1}(k-1)^{j-1}\binom{(k-1)(i-j)}{i-j}.$$

□

この補題を用いて，$N'_{(k-1)i+a}(a,0)$ の大きさを考えるとき，上限の方は以下のように粗く評価できる．

$$N'_{(k-1)i+a}(a,0)\leq\binom{(k-1)i+a}{i}\leq\left(e\,\frac{(k-1)i+a}{i}\right)^i.$$

次に，$(k-1)i+a\,(0\leq i\leq m)$ 回目の試行において初めてオープンコピーの数が 0 になる確率を $P(X_1\geq 1,\cdots,X_{(k-1)i+a-1}\geq 1,\ X_{(k-1)i+a}=0)$ と記す．ここではコンフィグレーションの最初から考えるので，$a=k$ とする．以下では，この補題を利用して，$\sum_{i=0}^{m}P(X_1\geq 1,\cdots,X_{(k-1)i+k-1}\geq 1,X_{(k-1)i+k}=0)$ の大きさを評価する．以下の補題では，α を最初の設定より大きくして $2\leq\alpha\leq 3$ とするが，その証明においては補題 1 以降を用いないので影響はない．

補題 7

$n\to\infty$ のとき，$k\to\infty$ とする．$2\leq\alpha\leq 3$ とし，m は $m^2=o(k)$ を満たす正の整数とする．このとき，最初から $(k-1)m+k$ 回までのコンフィギュレーションにおいて，オープンコピーの数が 0 にならない確率は少なくとも $1-2e^{-2}$ である．

証明

$1\leq t\leq(k-1)m+k$ においてバックエッジの起こる事象を B とし，起こらない事象を $\overline{B}$ とする．前述の議論に従えば，最初から $j=(k-1)m+k$ 回までのコンフィギュレーションにおいて，$\overline{B}$ の条件下でオープンコピーの数が 0 になる確率は，$\sum_{i=0}^{m}P(S_1\geq 1,\cdots,S_{(k-1)i+k-1}\geq 1,S_{(k-1)i+k}=0\,|\,\overline{B})$ となる．ここで，i は次数 k の未開示な頂点のコピーを選択する回数である．また，当試行の任意の時点において，次数 k と次数 1 の頂点のコピーをそれぞれ選択する確率は高々 $\frac{kd_k(n)}{M-2j}$ および $\frac{d_1(n)}{M-2j}$ である．なお，表記を簡単にするため，前者を p とし，後者を q とする．このとき，最初の 3 つの項は，オープンコピーの数が 0 になる回数を近似せずに以下のように求める．

$$P(S_1 \geq 1, \cdots, S_{k-1} \geq 1, S_k = 0 \,|\, \overline{B}) \leq q^k,$$

$$P(S_1 \geq 1, \cdots, S_{2k-2} \geq 1, S_{2k-1} = 0 \,|\, \overline{B}) \leq kpq^{2k-2},$$

$$P(S_1 \geq 1, \cdots, S_{3k-3} \geq 1, S_{3k-2} = 0 \,|\, \overline{B}) \leq \frac{3k(k-1)}{2} p^2 q^{3k-4}.$$

次に，第 4 項以降を考えるために，補題 6 から以下の不等式を得る．なお，2 番目の不等式は $i \geq 3$ から導かれる．

$$N_{(k-1)i+k}(k, 0) = \binom{(k-1)i+k}{i} \leq \left(e\,\frac{(k-1)i+k}{i}\right)^i \leq \left(\frac{3e(k-1)}{2}\right)^i.$$

よって，

$$\begin{aligned}
&\sum_{i=0}^{m} P(S_1 \geq 1, \cdots, S_{(k-1)i+k-1} \geq 1, S_{(k-1)i+k} = 0 \,|\, \overline{B}) \\
&\leq q^k + kpq^{2k-2} + \frac{3k(k-1)}{2} p^2 q^{3k-4} + \sum_{i=3}^{m} \binom{(k-1)i+k}{i} p^i q^{(k-2)i+k} \\
&\leq q^k + kpq^{2k-2} + \frac{3k(k-1)}{2} p^2 q^{3k-4} + \sum_{i=3}^{m} \left(\frac{3e(k-1)}{2}\right)^i p^i q^{(k-2)i+k} \\
&\leq q^k + kpq^{2k-2} + \frac{3k(k-1)}{2} p^2 q^{3k-4} + \frac{\left(\frac{3e(k-1)}{2}\right)^3 p^3 q^{4k-6}}{1 - \frac{3e(k-1)}{2} pq^{k-2}} \leq \frac{2}{e^2}
\end{aligned}$$

なお，最後から 2 番目および最後の不等式は以下の評価による．

$$p = \frac{kd_k(n)}{M-2j} = \frac{kd_k(n)}{M}\left(1 - \frac{2j}{M}\right)^{-1} = \frac{(\alpha+1)k}{(k-1)(k+\alpha)}\left(1 + O\left(\frac{km}{n}\right)\right),$$

$$\begin{aligned}
q^k &= \left(\frac{d_1(n)}{M-2j}\right)^k = \left(\frac{d_1(n)}{M}\right)^k \left(1 - \frac{2j}{M}\right)^{-k} \\
&= \left(\frac{k-2}{k-1}\,\frac{k}{k+\alpha}\left(1 - \frac{\alpha}{k(k-2)}\right)\right)^k \left(1 - \frac{2j}{M}\right)^{-k} \\
&= \left(\frac{k-2}{k-1}\,\frac{k}{k+\alpha}\right)^k \left(1 - \frac{\alpha}{k(k-2)}\right)^k \left(1 - \frac{2j}{M}\right)^{-k} \\
&\leq \left(1 - \frac{1}{k-1}\right)^k \left(1 - \frac{\alpha}{k+\alpha}\right)^k \left(1 - \frac{\alpha}{k(k-2)}\right)^k \left(1 - \frac{2j}{M}\right)^{-k} \\
&\leq e^{-(\alpha+1)}\left(1 + O\left(\frac{1}{k}\right)\right),
\end{aligned}$$

$$0 < \frac{3e(k-1)}{2} pq^{k-2} < 1.$$

よって，補題 1 を用いて以下の結果を得る．

$$\sum_{i=0}^{m} P(S_1 \geq 1, \cdots, S_{(k-1)i+k-1} \geq 1, S_{(k-1)i+k} = 0)$$
$$= P(B) \sum_{i=0}^{m} P(S_1 \geq 1, \cdots, S_{(k-1)i+k-1} \geq 1, S_{(k-1)i+k} = 0 \,|\, B)$$
$$+ P(\overline{B}) \sum_{i=0}^{m} P(S_1 \geq 1, \cdots, S_{(k-1)i+k-1} \geq 1, S_{(k-1)i+k} = 0 \,|\, \overline{B}) \leq o(1) + \frac{2}{e^2}.$$

この結果から，$n \to \infty$のとき，$(k-1)m + k$ 回までのコンフィギュレーションにおいて，オープンコピーの数が 0 にならない確率は少なくとも $1 - 2e^{-2}$ であることが示される． □

5. 課題

まず，次数 k の大きさが 3 の場合については，定理 1 と 2 の妥当性は証明されてないので，これを考察しなければならない．次に，$n \to \infty$のとき，$k \to \infty$となる場合について，ランダムグラフの大きさに関して提示した予想を検証する必要がある．

参考文献

Bollabás, B, *Random Graphs*, Second Edition, Cambridge Universitiy Press (2001).

Frieze, A and Karoňski, M, *Introduction to random graphs,* Cambridge University Press (2016).

Molloy, M and Reed, B, "A Critical Point for Random Graphs with a Given Degree Sequence", *Random Structures and Algorithms* 6, 161–180 (1995).

Molloy, M and Reed, B, "The Size of the Largest Component of a Random Graph on a Fixed Degree Sequence", *Combinatorics, Probability and Computing* 7, 295–306 (1998).

Molloy, M. and Reed, B, "A Critical Point for Random Graphs with a Given Degree Sequence", *citeseerx. ist. psu. edu* (2000).

浜口幸弘，確率的に生成されるネットワークの規模の評価Ⅰ—特定の次数列を満たす大規模ランダムグラフについて—，明治学院大学経済研究 147 (2014).

浜口幸弘，確率的に生成されるネットワークの規模の評価Ⅲ—特定の次数列を満たす大規模ランダムグラフについて—，明治学院大学経済研究 149 (2016).