A quantitative version of the Blow-up Lemma

by

Gábor N. Sárközy

Computer Science
Technical Report
Series

WORCESTER POLYTECHNIC INSTITUTE

Computer Science Department
100 Institute Road, Worcester, Massachusetts 01609-2280
A quantitative version of the Blow-up Lemma

Gábor N. Sárközy*
Alfréd Rényi Institute of Mathematics
Hungarian Academy of Sciences
Budapest, P.O. Box 127
Budapest, Hungary, H-1364

and

Computer Science Department
Worcester Polytechnic Institute
Worcester, MA, USA 01609

gsarkozy@cs.wpi.edu

Abstract

In this paper we give a quantitative version of the Blow-up Lemma.

1 Introduction

1.1 Notations and definitions

All graphs are simple, that is, they have no loops or multiple edges. $v(G)$ is the number of vertices in $G$ (order), $e(G)$ is the number of edges in $G$ (size). $\deg(v)$ (or $\deg_G(v)$) is the degree of vertex $v$ (within the graph $G$), and $\deg(v, Y)$ (or $\deg_G(v, Y)$) is the number of neighbors of $v$ in $Y$. $\delta(G)$ and $\Delta(G)$ are the minimum degree and the maximum degree of $G$. $N(x)$ (or $N_G(x)$) is the set of neighbors of the vertex $x$, and $e(X, Y)$ is the number of edges between $X$ and $Y$. A bipartite graph $G$ with color-classes $A$ and $B$ and edges $E$ will sometimes be written as $G = (A, B, E)$. For disjoint $X, Y$, we define the density

$$d(X, Y) = \frac{e(X, Y)}{|X| \cdot |Y|}.$$  

*Research supported in part by OTKA Grant No. K104343.
The density of a bipartite graph $G = (A, B, E)$ is the number
\[ d(G) = d(A, B) = \frac{|E|}{|A| \cdot |B|}. \]
For two disjoint subsets $A, B$ of $V(G)$, the bipartite graph with vertex set $A \cup B$ which has all the edges of $G$ with one endpoint in $A$ and the other in $B$ is called the pair $(A, B)$.
A pair $(A, B)$ is $\varepsilon$-regular if for every $X \subset A$ and $Y \subset B$ satisfying
\[ |X| > \varepsilon|A| \quad \text{and} \quad |Y| > \varepsilon|B| \]
we have
\[ |d(X, Y) - d(A, B)| < \varepsilon. \]
A pair $(A, B)$ is $(\varepsilon, d, \delta)$-super-regular if it is $\varepsilon$-regular with density at least $d$ and furthermore,
\[ \deg(a) \geq \delta|B| \quad \text{for all} \quad a \in A, \]
\[ \text{and} \quad \deg(b) \geq \delta|A| \quad \text{for all} \quad b \in B. \]
$H$ is embeddable into $G$ if $G$ has a subgraph isomorphic to $H$, that is, if there is a one-to-one map (injection) $\varphi : V(H) \to V(G)$ such that $\{x, y\} \in E(H)$ implies $\{\varphi(x), \varphi(y)\} \in E(G)$.

1.2 A quantitative version of the Blow-up Lemma

The Blow-up Lemma [8, 9] has been a successful tool in extremal graph theory. There are now at least four new proofs for the Blow-up Lemma since the original appeared; an algorithmic proof [9], a hypergraph-packing approach [17], a proof based on counting perfect matchings in (Szemerédi-) regular graphs [16], and its constructive version in [18]. Very recently the Blow-up Lemma has been generalized to hypergraphs by Keevash [6] and to $d$-arrangeable graphs by Böttcher, Kohayakawa, Taraz, and Würfl [3]. The Blow-up Lemma has been cited in close to 100 papers (see e.g. [1, 4, 7, 10, 11, 12, 13, 15, 16, 17, 18]). See also the discussion on the Regularity Lemma and the Blow-up Lemma on pages 803-804 in the Handbook of Graph Theory [2] or the survey paper [14].
In either of our proofs [8, 9], the dependence of the parameters was not computed explicitly (in particular the dependence of $\varepsilon$ on the other parameters). In this paper we give a quantitative version, i.e. we compute explicitly the parameters.

Theorem 1 (A quantitative version of the Blow-up Lemma). There exists an absolute constant $C$ such that, given a graph $R$ of order $r \geq 2$ and positive parameters
$d$, $\delta$, and $\Delta$, for any $0 < \epsilon < \left(\frac{d \alpha}{\delta \Delta} \right)^C$ the following holds. Let $N$ be an arbitrary positive integer, and let us replace the vertices of $R$ with pairwise disjoint $N$-sets $V_1, V_2, \ldots, V_r$ (blowing up). We construct two graphs on the same vertex-set $V = \bigcup V_i$. The graph $R(N)$ is obtained by replacing all edges of $R$ with copies of the complete bipartite graph $K_{N,N}$, and a sparser graph $G$ is constructed by replacing the edges of $R$ with some $(\epsilon, d, \delta)$-super-regular pairs. If a graph $H$ with $\Delta(H) \leq \Delta$ is embeddable into $R(N)$ then it is already embeddable into $G$.

Our proof is almost identical to the proof in [9]. Of course one difference is that we have to compute explicitly the dependence between the parameters. Furthermore, this is also a slight strengthening of the original statement as there can be a small number of exceptional vertices which may have smaller degrees ($\delta$ may be much smaller than $d$). We note that the recent “arrangeable” Blow-up Lemma [3] is also quantitative, but first of all the bound on $\epsilon$ is somewhat weaker and second it does not allow for the strengthening mentioned above. However, in a recent application [5] we needed precisely this strengthening. We believe that this quantitative version of the Blow-up Lemma will find other applications as well.

In Section 2 we give the embedding algorithm. In Section 3 we show that the algorithm is correct.

2 The algorithm

The main idea of the algorithm is the following. We embed the vertices of $H$ one-by-one by following a greedy algorithm, which works smoothly until there is only a small proportion of $H$ left, and then it may get stuck hopelessly. To avoid that, we will set aside a positive proportion of the vertices of $H$ as buffer vertices. Most of these buffer vertices will be embedded only at the very end by using a König-Hall argument.

2.1 Preprocessing

We will assume that $|V(H)| = |V(G)| = |\bigcup V_i| = n = rN$. We will assume for simplicity, that the density of every super-regular pair in $G$ is exactly $d$. This is not a significant restriction, otherwise we just have to put everywhere the actual density instead of $d$.

We will use the following parameters:

$$\epsilon \ll \epsilon' \ll \epsilon'' \ll \epsilon''' \ll d''' \ll d'' \ll d' \ll d,$$

where $a \ll b$ means that $a$ is small enough compared to $b$. For example we can select
the parameters in the following explicit way:

\[ d' = \frac{\delta d^3}{8r \Delta}, \quad d'' = (d')^3 = \left( \frac{\delta d^3}{8r \Delta} \right)^3, \quad d''' = (d'')^2 = \left( \frac{\delta d^6}{8r^2 \Delta} \right)^2, \quad \varepsilon''' = (d''')^2 = \left( \frac{\delta d^{12}}{8r \Delta} \right)^2, \]

\[ \varepsilon'' = (\delta d^2) \left( \frac{\delta d^2}{8r \Delta} \right)^2 (\varepsilon''')^3 = \left( \frac{\delta d^{50}}{8r \Delta} \right), \quad \varepsilon' = (\delta d^2) \left( \frac{\delta d^2}{8r \Delta} \right)^2 (\varepsilon''')^3 = \left( \frac{\delta d^{164}}{8r \Delta} \right) \quad \text{and} \]

\[ \varepsilon = (\varepsilon')^2 = \left( \frac{\delta d^2}{8r \Delta} \right)^{328}. \]

For easier reading, we will mostly use the letter \( x \) for vertices of \( H \), and the letter \( v \) for vertices of the host graph \( G \).

Given an embedding of \( H \) into \( R(N) \), it defines an assignment

\[ \psi : V(H) \rightarrow \{ V_1, V_2, \ldots, V_r \}, \]

and we want to find an embedding

\[ \varphi : V(H) \rightarrow V(G), \quad \varphi \text{ is one-to-one} \]

such that \( \varphi(x) \in \psi(x) \) for all \( x \in V(H) \). We will write \( X_i = \psi^{-1}(V_i) \) for \( i = 1, 2, \ldots, r \).

Before we start the algorithm, we order the vertices of \( H \) into a sequence \( S = (x_1, x_2, \ldots, x_n) \) which is more or less, but not exactly, the order in which the vertices will be embedded (certain exceptional vertices will be brought forward). Let \( m = rd^r N \). For each \( i \), choose a set \( B_i \) of \( d^r N \) vertices in \( X_i \) such that any two of these vertices are at a distance at least 3 in \( H \). (This is possible, for \( H \) is a bounded degree graph.) These vertices \( b_1, \ldots, b_m \) will be called the buffer vertices and they will be the last vertices in \( S \).

The order \( S \) starts with the neighborhoods \( N_H(b_1), N_H(b_2), \ldots, N_H(b_m) \). The length of this initial segment of \( S \) will be denoted by \( T_0 \). Thus \( T_0 = \sum_{i=1}^{m} |N_H(b_i)| \leq \Delta m \).

The rest of \( S \) is an arbitrary ordering of the leftover vertices of \( H \).

### 2.2 Sketch of the algorithm

In Phase 1 of the algorithm we will embed the vertices in \( S \) one-by-one into \( G \) until all non-buffer vertices are embedded. For each \( x_j \) not embedded yet (including the buffer vertices) we keep track of an ever shrinking host set \( H_{t,x_j} \), that \( x_j \) is confined to at time \( t \), and we only make a final choice for the location of \( x_j \) from \( H_{t,x_j} \) at time \( j \). At time 0, \( H_{0,x_j} \) is the cluster that \( x_j \) is assigned to. For technical reasons we will also maintain another similar set, \( C_{t,x_j} \), where we will ignore the possibility that some
vertices are occupied already. $Z_t$ will denote the set of occupied vertices. Finally we will maintain a set $\text{Bad}_t$ of exceptional pairs of vertices.

In Phase 2, we embed the leftover vertices by using a König-Hall type argument.

### 2.3 Embedding Algorithm

At time 0, set $C_{0,x} = H_{0,x} = \psi(x)$ for all $x \in V(H)$. Put $T_1 = d'n$.

**Phase 1.**

For $t \geq 1$, repeat the following steps.

**Step 1** (Extending the embedding). We embed $x_t$. Consider the vertices in $H_{t-1,x_t}$. We will pick one of these vertices as the image $\varphi(x_t)$ by using the Selection Algorithm (described below in Section 2.4).

**Step 2** (Updating). We set $Z_t = Z_{t-1} \cup \{\varphi(x_t)\}$, and for each unembedded vertex $y$ (i.e. the set of vertices $x_j, t < j \leq n$), set

$$C_{t,y} = \begin{cases} 
C_{t-1,y} \cap N_G(\varphi(x_t)) & \text{if } \{x_t, y\} \in E(H) \\
C_{t-1,y} & \text{otherwise},
\end{cases}$$

and $H_{t,y} = C_{t,y} \setminus Z_t$.

We do not change the ordering at this step.

**Step 3** (Exceptional vertices in $G$).

1. If $t \notin \{1, T_0\}$, then go to Step 4.
2. If $t = 1$, then we do the following (this is the part that is new compared to the proof in [9]). We find the 1st exceptional set (denoted by $E_1$) consisting of those exceptional vertices $v \in V_i, 1 \leq i \leq r$ for which there exists a $j \neq i$ such that $(V_i, V_j)$ is $(\varepsilon, d, \delta)$-super-regular, yet

$$d_{G}(v, V_j) < (d - \varepsilon)|V_j|.$$  

(Note that $d_{G}(v, V_j) \geq \delta|V_j|$ always holds by super-regularity.) $\varepsilon$-regularity implies that $|E_1| \leq r\varepsilon N$. We are going to change slightly the order of the vertices in $S$. We choose a set $E_1^H$ of nonbuffer vertices $x \in H$ of size $\Sigma_{i=1}^r |E_1^i|$ (more precisely $|E_1^i|$ vertices from $X_i$ for all $1 \leq i \leq r$) such that they are at a distance at
least 3 from each other. This is possible since $H$ is a bounded degree graph and $\sum_{i=1}^{r} |E^1_i|$ is very small. We bring the vertices in $E^1_H$ forward, followed by the remaining vertices in the same relative order as before. For simplicity we keep the notation $(x_1, x_2, \ldots, x_n)$ for the resulting order. Furthermore, we slightly change the value of $T_0$ to $T_0 = |E^1_H| + \sum_{i=1}^{m} |N_H(b_i)|$.

3. If $t = T_0$, then we do the following. We find the 2nd exceptional set (denoted by $E^2_i$) consisting of those exceptional vertices $v \in V_i$, $1 \leq i \leq r$ for which $v$ is not covered yet in the embedding and

$$|\{b : b \in B_i, v \in C_{t,b}\}| < d'' |B_i|.$$ 

Once again we are going to change slightly the order of the remaining unembedded vertices in $S$. We choose a set $E^2_H$ of unembedded nonbuffer vertices $x \in H$ of size $\sum_{i=1}^{r} |E^2_i|$ (more precisely $|E^2_i|$ vertices from $X_i$ for all $1 \leq i \leq r$) with

$$H_{t,x} = H_{0,x} \setminus \{\varphi(x_j) : j \leq t\} = \psi(x) \setminus \{\varphi(x_j) : j \leq t\}.$$ 

Thus in particular, if $x \in X_i$, then $E^2_i \subset H_{t,x}$. Again we may choose the vertices in $E^2_H$ as vertices in $H$ that are at a distance at least 3 from each other and any of the vertices embedded so far. We are going to show later in the proof of correctness that this is possible since $H$ is a bounded degree graph and $\sum_{i=1}^{r} |E^2_i|$ is very small as well. We bring the vertices in $E^2_H$ forward, followed by the remaining unembedded vertices in the same relative order as before. Again, for simplicity we keep the notation $(x_1, x_2, \ldots, x_n)$ for the resulting order.

**Step 4** (Exceptional vertices in $H$).

1. If $T_1$ does not divide $t$, then go to Step 5.

2. If $T_1$ divides $t$, then we do the following. We find all exceptional unembedded vertices $y \in H$ such that $|H_{t,y}| \leq (d')^2n$. Once again we slightly change the order of the remaining unembedded vertices. We bring these exceptional vertices forward (even if they are buffer vertices), followed by the non-exceptional vertices in the same relative order as before. Again for simplicity we still use the notation $(x_1, x_2, \ldots, x_n)$ for the new order. Note that it will follow from the proof, that if $t \leq 2T_0$, then we do not find any exceptional vertices in $H$, so we do not change the ordering at this step.

**Step 5** - If there are no more unembedded non-buffer vertices left, then set $T = t$ and go to Phase 2, otherwise set $t \leftarrow t + 1$ and go back to Step 1.

**Phase 2**

Find a system of distinct representatives of the sets $H_{T,y}$ for all unembedded $y$ (i.e. the set of vertices $x_j$, $T < j \leq n$).
2.4 Selection Algorithm

We distinguish two cases. Let $E_H = E^1_H \cup E^2_H$.

**Case 1.** $x_t \notin E_H$.

We choose a vertex $v \in H_{t-1}$ as the image $\varphi(x_t)$ for which the following hold for all unembedded $y$ with $\{x_t, y\} \in E(H)$,

$$
(d - \varepsilon)|H_{t-1,y}| \leq deg_G(v, H_{t-1,y}) \leq (d + \varepsilon)|H_{t-1,y}|, \tag{1}
$$

$$
(d - \varepsilon)|C_{t-1,y}| \leq deg_G(v, C_{t-1,y}) \leq (d + \varepsilon)|C_{t-1,y}| \tag{2}
$$

and

$$
(d - \varepsilon)|C_{t-1,y} \cap C_{t-1,y'}| \leq deg_G(v, C_{t-1,y} \cap C_{t-1,y'}) \leq (d + \varepsilon)|C_{t-1,y} \cap C_{t-1,y'}|, \tag{3}
$$

for at least a $(1 - \varepsilon')$ proportion of the unembedded vertices $y'$ with $\psi(y') = \psi(y)$ and $\{y, y'\} \notin \text{Bad}_{t-1}$. Then we get $\text{Bad}_t$ by taking the union of $\text{Bad}_{t-1}$ and the set of all of those pairs $\{y, y'\}$ for which (3) does not hold for $v = \varphi(x_t)$, $C_{t-1,y}$ and $C_{t-1,y'}$. Thus note that we add at most $\Delta \varepsilon' N$ new pairs to $\text{Bad}_t$.

**Case 2.** $x_t \in E_H$.

If $x_t \in X_i \cap E^1_H, l = 1, 2$, then we choose an arbitrary vertex of $E^l_i$ as $\varphi(x_t)$. Note that for all $y \in N_H(x_t)$, we have $C_{t-1,y} = \psi(y)$,

$$
deg_G(\varphi(x_t), C_{t-1,y}) = deg_G(\varphi(x_t), \psi(y)) \geq \delta N = \delta |C_{t-1,y}|, \tag{4}
$$

and

$$
\deg_G(\varphi(x_t), H_{t-1,y}) \geq \deg_G(\varphi(x_t), \psi(y)) - T_0 - |E_H| \geq \delta N - 2\Delta rd'N \geq \frac{\delta}{2}N \tag{5}
$$

(using our choice of parameters). Here we used super-regularity and the fact that $|E_H| \ll \Delta m$ which will be shown later (Lemma 3).

3 Proof of correctness

The following claims state that our algorithm finds a good embedding of $H$ into $G$.

**Claim 1.** Phase 1 always succeeds.

**Claim 2.** Phase 2 always succeeds.

If at time $t$, $S$ is a set of unembedded vertices $x \in H$ with $\psi(x) = V_i$ (here and throughout the proof when we talk about time $t$, we mean after Phase 1 is executed
for time \( t \), so for example \( x_t \) is considered embedded at time \( t \), then we define the bipartite graph \( U_t \) as follows. One color class is \( S \), the other is \( V_i \), and we have an edge between an \( x \in S \) and a \( v \in V_i \) whenever \( v \in C_{t,x} \).

In the proofs of the above claims the following lemma will play a major role. First we prove the lemma for \( t \leq T_0 \), from this we deduce that \( |E_H| \) is small, then we prove the lemma for \( T_0 < t \leq T \).

**Lemma 2.** We are given integers \( 1 \leq i \leq r \), \( 1 \leq t \leq T_0 \) and a set \( S \subset X_i \) of unembedded vertices at time \( t \) with \( |S| \geq (d'')^2|X_i| = (d'')^2N \). If we assume that Phase 1 succeeded for all time \( t' \) with \( t' \leq t \), then apart from an exceptional set \( F \) of size at most \( \varepsilon''N \), for every vertex \( v \in V_i \) we have the following

\[
\deg_{U_t}(v) = |\{ x : x \in S, v \in C_{t,x} \}| \geq (1 - \varepsilon'')d(U_t)|S| \geq \frac{d\Delta}{2}|S|.
\]

**Proof.** In the proof of this lemma we will use the “defect form” of the Cauchy-Schwarz inequality (just as in the original proof of the Regularity Lemma [19]): if

\[
\sum_{k=1}^{m} X_k = \frac{m}{n} \sum_{k=1}^{n} X_k + D \quad (m \leq n)
\]

then

\[
\sum_{k=1}^{n} X_k^2 \geq \frac{1}{n} \left( \sum_{k=1}^{n} X_k \right)^2 + \frac{D^2n}{m(n - m)}.
\]

Assume indirectly that the statement in Lemma 2 is not true, that is, \( |F| > \varepsilon''N \). We take an \( F_0 \subset F \) with \( |F_0| = \varepsilon''N \). Let us write \( \nu(t, x) \) for the number of neighbors (in \( H \)) of \( x \) embedded by time \( t \). Then in \( U_t \) using the left side of (2) we get

\[
e(U_t) = d(U_t)|S||V_i| = \sum_{v \in V_i} \deg_{U_t}(v) = \sum_{x \in S} \deg_{U_t}(x)
\]

\[
= \sum_{x \in S} |C_{t,x}| \geq \sum_{x \in S} (d - \varepsilon)\nu(t, x)N - \Delta r^2 \varepsilon N^2 \geq (d - \varepsilon)\Delta |S|N - \Delta r^2 \varepsilon N^2 \geq \frac{d\Delta}{2}|S|N, \quad (6)
\]

where the error term comes from the neighbors of elements of \( E^1_H \) (we are yet to start the embedding of the vertices in \( E^2_H \)), since for them we cannot guarantee the same lower bound.

We also have

\[
\sum_{x \in S} \sum_{x' \in S} |N_{U_t}(x) \cap N_{U_t}(x')| = \sum_{x \in S} \sum_{x' \in S} |C_{t,x} \cap C_{t,x'}|
\]

\[
\leq \sum_{x \in S} \sum_{x' \in S} (d + \varepsilon)\nu(t, x)N + |S|N + \Delta^2 |S|N + 2\Delta r^2 \varepsilon |S|N^2 + 2\Delta \varepsilon' N^3
\]
\[ \leq \sum_{x \in S} \sum_{x' \in S} (d + \varepsilon)^{\nu(t,x) + \nu(t,x')} N + 5\Delta \varepsilon' N^3. \quad (7) \]

The error terms come from the following \((x, x')\) pairs. For each such pair we estimate \(|C_{t,x} \cap C_{t,x'}| \leq N\). The first error term comes from the pairs where \(x = x'\). The second error term comes from those pairs \((x, x')\) for which \(N_H(x) \cap N_H(x') \neq \emptyset\). The number of these pairs is at most \(|S|\Delta(\Delta - 1) \leq \Delta^2|S|\). The third error term comes from those pairs \((x, x')\) for which \(x\) or \(x'\) is a neighbor of an element of \(E^*_{H}\). Finally we have the pairs for which \(\{x, x'\} \in \text{Bad}_t\). The number of these pairs is at most \(2t\Delta \varepsilon' N \leq 2\Delta \varepsilon' N^2\).

Next we will use the Cauchy-Schwarz inequality with \(m = \varepsilon'' N\) and the variables \(X_k, k = 1, \ldots, N\) are going to correspond to \(\text{deg}_{U_i}(v), v \in V_i\) (and the first \(m\) variables to degrees in \(F_0\)). Then we have

\[ |D| = \varepsilon'' \sum_{v \in V_i} \text{deg}_{U_i}(v) - \sum_{v \in F_0} \text{deg}_{U_i}(v) \]
\[ \geq \varepsilon'' \sum_{v \in V_i} \text{deg}_{U_i}(v) - \varepsilon''(1 - \varepsilon'')d(U_i)|S|N = (\varepsilon'')^2d(U_i)|S|N. \quad (8) \]

Then using (6), (8) and the Cauchy-Schwarz inequality we get

\[ \sum_{x \in S} \sum_{x' \in S} |N_{U_i}(x) \cap N_{U_i}(x')| = \sum_{v \in V_i} (\text{deg}_{U_i}(v))^2 \]
\[ \geq \frac{1}{N} \left( \sum_{v \in V_i} \text{deg}_{U_i}(v) \right)^2 + (\varepsilon'')^3d(U_i)^2N|S|^2 \]
\[ \geq \frac{1}{N} \left( \sum_{x \in S} (d - \varepsilon)^{\nu(t,x)} N - \Delta r^2 N^2 \right)^2 + (\varepsilon'')^3d(U_i)^2N|S|^2 \]
\[ \geq \sum_{x \in S} \sum_{x' \in S} (d - \varepsilon)^{\nu(t,x) + \nu(t,x')} N - 2\Delta \varepsilon' N^3 + (\varepsilon'')^3(d - \varepsilon)^{2\Delta} N|S|^2, \]

which is a contradiction with (7), since \(|S| \geq (d^m)^2 N\),

\[ ((d + \varepsilon)^{\nu(t,x) + \nu(t,x')} - (d - \varepsilon)^{\nu(t,x) + \nu(t,x')}) \ll \Delta \varepsilon, \]

and

\[ (\varepsilon'')^3(d - \varepsilon)^{2\Delta}(d^m)^2 \gg \frac{d^{2\Delta}}{2}(d^m)^2(\varepsilon'')^3 \geq \Delta \varepsilon' \gg \Delta \varepsilon, \]

by the choice of the parameters.

An easy consequence of Lemma 2 is the following lemma.

**Lemma 3.** In Step 3 we have \(|E_{1}^2| \leq \varepsilon'' N\) for every \(1 \leq i \leq r\).
\textbf{Proof.} Indeed applying Lemma 2 with \( t = T_0 \) and \( S = B_i \) (so we have \(|S| = |B_i| = d'N > (d'')^2N\)) we get

\[(1 - \varepsilon'')d(U_t)|S| \geq \frac{d\Delta}{2}|S| > d''|S|,
\]

and \( E_i^2 \subset F \).

From this we can prove Lemma 2 for \( t > T_0 \) with \( \varepsilon''' \) instead of \( \varepsilon'' \).

\textbf{Lemma 4.} We are given integers \( 1 \leq i \leq r \), \( T_0 < t \leq T \) and a set \( S \subset X_i \) of unembedded vertices at time \( t \) with \(|S| \geq (d'')^2|X_i| = (d'')^2N \). If we assume that Phase 1 succeeded for all time \( t' \) with \( t' \leq t \), then apart from an exceptional set \( F \) of size at most \( \varepsilon'''N \), for every vertex \( v \in V_i \) we have the following

\[ \deg_{U_t}(v) = |\{x : x \in S, v \in C_{t,x}\}| \geq (1 - \varepsilon''')d(U_t)|S| \geq \frac{d\Delta}{2}|S| \text{.} \]

\textbf{Proof.} We only have to pay attention to the neighbors of the elements of \( E_i^2 \), otherwise the proof is the same as the proof of Lemma 2 with \( \varepsilon''' \) instead of \( \varepsilon'' \). In (6) the error term becomes \( \Delta r\varepsilon''N^2 \), coming from the neighbors of elements of \( E_i^2 \).

In (7) we have more bad pairs, namely all pairs \((x, x')\) where \( x \) or \( x' \) is a neighbor of an element of \( E_i^2 \). These give an additional error term of \( 2\Delta r\varepsilon'''|S|N^2 \). However, the contradiction still holds, since

\[(\varepsilon''')^3(d - \varepsilon)^2(d'')^2 \gg \frac{d\Delta}{2}(d'')^2(\varepsilon''')^3 \geq \Delta\varepsilon'',\]

by the choice of the parameters.

An easy consequence of Lemmas 2 and 4 is the following lemma.

\textbf{Lemma 5.} We are given integers \( 1 \leq i \leq r \), \( 1 \leq t \leq T \), a set \( S \subset X_i \) of unembedded vertices at time \( t \) with \(|S| \geq d''|X_i| = d''N \) and a set \( A \subset V_i \) with \(|A| \geq d''|V_i| = d''N \). If we assume that Phase 1 succeeded for all time \( t' \) with \( t' \leq t \), then apart from an exceptional set \( S' \) of size at most \( (d'')^2N \), for every vertex \( x \in S \) we have the following

\[ |A \cap C_{t,x}| \geq \frac{|A|}{2N}|C_{t,x}|. \]  

\textbf{Proof.} Assume indirectly that the statement is not true, i.e. there exists a set \( S' \subset S \) with \(|S'| > (d'')^2N \) such that for every \( x \in S' \) (9) does not hold. Once again we consider the bipartite graph \( U_t = U_t(S', V_i) \). We have

\[ \sum_{x \in S'} \deg_{U_t}(v) = \sum_{x \in S'} |A \cap C_{t,x}| < \frac{|A|}{2N} \sum_{x \in S'} |C_{t,x}| = \frac{|A|}{2N} d(U_t)|S'|N. \]
On the other hand, applying Lemmas 2 or 4 for $S'$ we get
$$\sum_{v \in A} \text{deg}_{U_t}(v) \geq (1 - \varepsilon'')d(U_t)|S'|(|A| - \varepsilon''N)$$
contradicting the previous inequality.

Finally we have

**Lemma 6.** For every $1 \leq t \leq T$ and for every vertex $y$ that is unembedded at time $t$, if we assume that Phase 1 succeeded for all time $t'$ with $t' \leq t$, then we have the following at time $t$
$$|H_{t,y}| > d''N. \quad (10)$$

**Proof.** We apply Lemma 5 with $S_t$ the set of all unembedded vertices in $X_i$ at time $t$, and $A_t = V_i \setminus Z_t$ (all uncovered vertices). Then for all but at most $(d''')^2N$ vertices $x \in S_t$ using (2) and (4) we get
$$|H_{t,x}| = |A_t \cap C_{t,x}| \geq |A_t| |C_{t,x}| \geq \frac{d'}{2N} \delta(d - \varepsilon)^4 N \geq (d')^2 N, \quad (11)$$
if $|A_t| \geq (d'/2)N$. We will show next that in fact for $1 \leq t \leq T$, we have
$$|A_t| \geq |A_T| \geq (d' - d'')N \quad \left(\geq \frac{d'}{2N}\right),$$
so (11) always holds. Assume indirectly that this is not the case, i.e. there exists a $1 \leq T' < T$ for which,
$$|A_{T'}| \geq (d' - d'')N \quad \text{but} \quad |A_{T'+1}| < (d' - d'')N.$$

From the above at any given time $t$ for which $T_1|t$ and $1 \leq t \leq T'$, in Step 4 we find at most $(d'')^2N$ exceptional vertices in $X_i$. Hence, altogether we find at most
$$\frac{1}{d''}(d'')^2N \ll d''N$$
exceptional vertices in $X_i$ up to time $T'$. However, this implies that at time $T'$ we still have many more than $(d' - d'')N$ unembedded buffer vertices in $X_i$, which in turn implies that $|A_{T'+1}| \gg (d' - d'')N$, a contradiction. Thus we have
$$|A_T| \geq (d' - d'')N, \quad T \leq rN - rd'N + rd''N,$$
at time $T$ (or in Phase 2) we have at least $(d' - d'')N$ unembedded buffer vertices in each $X_i$, and furthermore, for every $1 \leq t \leq T$ for all but at most $(d'')^2N$ vertices $x \in S_t$ we have
$$|H_{t,x}| > (d')^2 N.$$
Let us pick an arbitrary $1 \leq t \leq T$ and an unembedded $y$ at time $t$ (with $\psi(y) = V_i$). We have to show that (10) holds. Let $kd''n = kT_1 \leq t < (k + 1)T_1$ for some $0 \leq k \leq T/T_1$. We distinguish two cases:

**Case 1.** $y$ was not among the at most $(d'')^2N$ exceptional vertices of $X_i$ found in Step 4 at time $kT_1$. Then

$$|H_{t,y}| \geq \left(\frac{\delta}{2}(d - \varepsilon)^\Delta(d')^2 - rd''\right) N.$$  

Indeed, at time $kT_1$ we had $|H_{kT_1,y}| \geq (d')^2N$. Until time $t$, $H_{t,y}$ could have been cut at most once to a $\geq (\delta/2)$-fraction (if $y$ is a neighbor of an element of $E_H$, there can be at most one such $E_H$-neighbor) and at most $\Delta$ times to a $\geq (d - \varepsilon)$-fraction (using (1) and (5)), and precisely $t - kT_1 \leq T_1 = rd''N$ new vertices were covered.

**Case 2.** $y$ was among the at most $(d'')^2N$ exceptional vertices of $X_i$ found in Step 4 at time $kT_1$. Then

$$|H_{t,y}| \geq \left(\frac{\delta}{2}(d - \varepsilon)^\Delta(d')^2 - rd'' - r(d'')^2\right) N,$$

since at time $(k - 1)T_1$ (we certainly must have $k \geq 2$), $y$ was not exceptional, and because the exceptional vertices were brought forward we have $t \leq kT_1 + r(d'')^2N$. Thus in both cases we have $|H_{t,y}| > d''N$, as desired.

Finally we show that the selection algorithm always succeeds in selecting an image $\varphi(x_t)$.

**Lemma 7.** For every $1 \leq t \leq T$, if we assume that Phase 1 succeeded for all time $t'$ with $t' \leq t$, then Phase 1 succeeds for time $t$.

**Proof.** We only have to consider Case 1 in the selection algorithm. We choose a vertex $v \in H_{t-1,x_t}$ as the image $\varphi(x_t)$ which satisfies (1), (2) and (3). We have by Lemma 6,

$$|H_{t-1,x_t}| \geq d''N.$$  

By $\varepsilon$-regularity we have at most $2\varepsilon N$ vertices in $H_{t-1,x_t}$ which do not satisfy (1) and similarly for (2). For (3) we define an auxiliary bipartite graph $B$ as follows. One color class $W_1$ is the vertices in $H_{t-1,x_t}$ and the other class $W_2$ is the sets $C_{t-1,y} \cap C_{t-1,y'}$ for all pairs $\{y, y'\}$ where $\{x_t, y\} \in E(H)$, $\psi(y) = \psi(y')$, and $\{y, y'\} \notin \text{Bad}_{t-1}$. We put an edge between a $v \in W_1$ and an $S \in W_2$ if inequality (3) is not satisfied for $v$ and $S$. Let us assume indirectly that we have more than $\varepsilon' N$ vertices $v \in W_1$ with $\deg_B(v) > \varepsilon' |W_2|$. Then there must exist a $S \in W_2$ with

$$\deg_B(S) > \varepsilon' |W_1| \gg \varepsilon N.$$
However, this is a contradiction with $\varepsilon$-regularity since

$$|S| \geq (d - \varepsilon)2^\Delta N \gg \varepsilon N.$$ 

Here we used the fact that the pair corresponding to $S$ is not in $\text{Bad}_{t-1}$. Thus altogether we have at most $4\varepsilon N + \varepsilon'N \ll d''N$ vertices in $H_{t-1,x_t}$ that we cannot choose and thus the selection algorithm always succeeds in selecting an image $\varphi(x_t)$, proving Claim 1.

**Proof of Claim 2.** We want to show that we can find a system of distinct representatives of the sets $H_{T,x_j}, T < j \leq n$, where the sets $H_{T,x_j}$ belong to a given cluster $V_i$.

To simplify notation, let us denote by $Y$ the set of remaining vertices in $V_i$, and by $X$ the set of remaining unembedded (buffer) vertices assigned to $V_i$. If $x = x_j \in X$ then write $H_x$ for its possible location $H_{T,x_j}$ at time $T$. Also write $M = |X| = |Y|$. The König-Hall condition for the existence of a system of distinct representatives obviously follows from the following three conditions:

\begin{align*}
|H_x| > d''M & \quad \text{for all } x \in X, \quad (12) \\
|\bigcup_{x \in S} H_x| \geq (1 - d'')M & \quad \text{for all subsets } S \subset X, |S| \geq d''M, \quad (13) \\
|\bigcup_{x \in S} H_x| = M & \quad \text{for all subsets } S \subset X, |S| \geq (1 - d'')M. \quad (14)
\end{align*}

Equation (12) is an immediate consequence of Lemma 6, (13) is a consequence of Lemma 2. Finally to prove (14), we have to show that every vertex in $Y \subset V_i$ belongs to at least $d''M$ location sets $H_x$. However, this is trivial from the construction of the embedding algorithm, in Step 3 of Phase 1 we took care of the small number of exceptional vertices for which this is not true. This finishes the proof of Claim 2 and the proof of correctness.

**References**


