

**Lemma 5.3.** *There is an  $n_0$  such that the following holds for all  $n \geq n_0$ ,  $2^{-9} < \varepsilon < 1$ ,  $c > 0$ ,  $r, t \geq (\log n)^2$  and  $s \geq 20rt$ . Let  $G$  be an  $n$ -vertex  $(\varepsilon, c, s)$ -expander, let  $U \subseteq V(G)$  satisfy  $|U| \leq 2n/3$ . Then, in  $G$  we can find either*

- (a)  $\frac{|U|}{10r}$  pairwise vertex-disjoint stars of size  $t$ , whose centers are in  $U$  and whose leaves are in  $V(G) - U$ , or
- (b) a bipartite subgraph  $H$  with vertex classes  $U$  and  $X \subseteq V(G) - U$  such that
  - $|X| \geq \frac{\varepsilon|U|}{2(\log n)^c}$  and
  - every vertex in  $X$  has degree at least  $r$  in  $H$  and every vertex in  $U$  has degree at most  $2t$  in  $H$ .

**Proof.** Take a maximal collection  $\mathcal{C}$  of pairwise vertex-disjoint stars in  $G$  with  $t$  leaves, centres in  $U$  and leaves outside of  $U$ . Let  $C \subseteq U$  be the set of centres of these stars and  $L \subseteq V(G) - U$  be the set consisting of all their leaves. Suppose **a**) does not hold. Then we can assume that  $|C| \leq \frac{|U|}{10r}$  and thus  $|L| = |C| \cdot t \leq \frac{|U|}{10r} \cdot t$ , and, by the maximality of  $\mathcal{C}$ , that there is no vertex in  $U - C$  with at least  $t$  neighbours in  $G$  in  $V(G) - (U \cup L)$ . Thus,

$$|N_G(U - C)| \leq |C| + |L| + |U - C| \cdot t \leq \frac{|U|}{10r} + |C| \cdot t + |U - C| \cdot t < 2|U| \cdot t. \quad (6)$$

We now construct a set  $X \subseteq V(G) - U$  and a bipartite subgraph  $H$  with vertex classes  $U$  and  $X$  using the following process, starting with  $X_0 = \emptyset$  and setting  $H_0$  to be the graph with vertex set  $U \cup X_0$  and no edges. Let  $k = |V(G) - U|$  and label the vertices of  $V(G) - U$  arbitrarily as  $v_1, \dots, v_k$ . For each  $i \geq 1$ , if possible, pick a star  $S_i$  in  $G$  with centre  $v_i$  and  $r$  leaves in  $U$  such that the vertices in  $U$  in the graph  $H_{i-1} \cup S_i$  have degree at most  $2t$ , and let  $H_i = H_{i-1} \cup S_i$  and  $X_i = X_{i-1} \cup \{v_i\}$ , while otherwise we set  $H_i = H_{i-1}$  and  $X_i = X_{i-1}$ . Finally, let  $H = H_k$  and  $X = X_k = V(H_k) - U$ . We will now show that **b**) holds for this choice of  $H$  (with vertex classes  $U$  and  $X$ ).

Firstly, observe that every vertex of  $U$  has degree at most  $2t$  in  $H_i$  for each  $i \in [k]$  by construction, and that every vertex  $v_i$  in  $X$  has degree exactly  $r$  in  $H$ , so the second condition in **b**) holds. Thus, we only need to show that  $|X| \geq \frac{\varepsilon|U|}{2(\log n)^c}$  holds.

To see this, let  $U'$  be the set of vertices in  $U - C$  with degree exactly  $2t$  in  $H$ . As each vertex in  $U - C$  has fewer than  $t$  neighbours in  $G$  in  $X - L$  (due to the maximality of the collection of stars  $\mathcal{C}$ ), the vertices in  $U'$  must have at least  $t$  neighbours in  $H$  in  $X \cap L$ . As each vertex in  $X \cap L$  has  $r$  neighbours in  $H$ , we have

$$|U'| \leq \frac{r|X \cap L|}{t} \leq \frac{r}{t} \cdot |L| \leq \frac{r}{t} \cdot \frac{|U| \cdot t}{10r} = \frac{|U|}{10}.$$

Let  $B = C \cup U'$ , so that

$$|B| \leq \frac{|U|}{10r} + \frac{|U|}{10} \leq \frac{|U|}{2},$$

and, thus,  $|U - B| \geq \frac{|U|}{2}$ .

Then, by Proposition 5.2 applied to  $U - B$  with  $d = r$ , we have either  $|N_G(U - B)| \geq \frac{s|U - B|}{2r}$  or  $|N_{G,r}(U - B)| \geq \frac{\varepsilon|U - B|}{(\log n)^c}$ . As

$$\frac{s|U - B|}{2r} \geq \frac{s|U|}{4r} \geq 5t|U|,$$

the former inequality contradicts (6), so we have that  $|N_{G,r}(U - B)| \geq \frac{\varepsilon|U - B|}{(\log n)^c}$ . Every vertex  $v_i$  in  $N_{G,r}(U - B)$  has at least  $r$  neighbours in  $G$  in  $U - B$ , and vertices of  $U - B$  must all have degree strictly less than  $2t$  in  $H$  (as they are not in  $U'$ ). This implies that every  $v_i$  in  $N_{G,r}(U - B)$ , satisfies  $v_i \in X$ , since we could add it along with some  $r$  of its neighbours while constructing  $H$ . Hence,  $N_{G,r}(U - B) \subseteq X$ , and

$$|X| \geq |N_{G,r}(U - B)| \geq \frac{\varepsilon|U - B|}{(\log n)^c} \geq \frac{\varepsilon|U|}{2(\log n)^c},$$

as required. □