We want to prove the following lemma:

Key Lemma: For every $d>0$ and integer $l>0$, there are $½>e_{d,l}>0$, $N_{d,l}$, and $a_{d,l}>0$ such that if $X_1,\ldots,X_l$ are subsets of $V(G)$, each of size $N \geq N_{d,l}$ and for all $1\leq i<j\leq l$, the pair (X_i,X_j) is $e_{d,l}$-regular with density at least d then G contains at least $a_{d,l}N^l$ cliques each containing one vertex in each of the X_i.

Proof: It is enough to prove the result for $d<1/2$ as it then holds for all larger d as well. We proceed by induction on l. The base case when l is 2 is trivial, since by the definition of density, we can set $a_{d,2}$ to be d.

So we assume that l is at least three and that the lemma holds for (l-1,d) for all values of $d<1/2$.

We set $e_{d,l}$ to be the minimum of $(1/l, d/2. de_{d/2,l-1}/2)$. We set $N_{d,l}$ to be $2N_{d/2,l-1}/d$. We set $a_{d,l}$ to be $(d/2)^{l-1}(a_{d/2,l-1}/l)$.

For each i between 1 and l-1 we let Y_i be the set of vertices of X_l which have fewer than d/2 neighbours in X_i. Since $e_{d,l} \leq d/2$, the regularity of (X_i,X_l) implies that $|Y_i|$ is at most $e_{d,l}N \leq N/l$ and that the subset X^*_l of vertices of X_l in none of the Y_i has size at least N/l. We shall show that each vertex v in X^*_l is in $(d/2)^{l-1}a_{d/2,l-1}N^{l-1}$ cliques, each containing one vertex in each of the X_i. To this end, we let X^v_i be a subset of the neighbourhood of v in X_i whose size is the round up N^* of dN/2 and apply induction on $X^v_1,\ldots,X^v_{l-1}$.

Since $N_{d,l}$ is $2N_{d,l-1}/d$, N^* is at least $N_{d/2,l-1}$. So it remains to show that every pair (X^v_i,X^v_j) is $e_{d/2,l-1}$-regular. To this end, suppose, X'_i (resp. X'_j) is a subset of at least $e_{d/2,l-1}N^*$ vertices of X^v_i (resp. X^v_j). Then, since $N^*>dN/2$, the $e_{d,l}$-regularity of (X_i,X_j) implies that both $d(X'_i,X'_j)$ and $d(X^v_i,X^v_j)$ lie within $e_{d,l}$ of $d(X_i,X_j)$ so they differ by at most $2e_{d,l}<e_{d/2,l-1}$ and the proof is complete.

We immediately obtain the following:

Keyer Lemma: For every $d>0$ and integer $l>0$, there are $½> e_{d,l}>0$, $N_{d,l}$, and $a_{d,l}>0$ such that if H is a graph on vertices $v_1,\ldots,v_l$ and $X_1,\ldots,X_l$ are subsets of $V(G)$, each of size $N \geq N_{d,l}$ and for all $1\leq i<j\leq l$, the pair (X_i,X_j) is $e_{d,l}$-regular with density which is at least d if v_iv_j is an edge of G and at most 1-d otherwise then G contains at least $a_{d,l}N^{\wedge l}$ induced copies of H, each containing one vertex in each of the X_i.

Proof: A pair is e-regular in G if and only if it is e-regular in the complement of G, since the density of a pair in G and the complement of G sums to 1. We simply complement the edges between every pair corresponding to a non-edge of H and apply the Key Lemma.

We now present a variant of this result which allows us to require fewer sets at the expense of a double-sided density bound.

Definition: t(H) is the minimum integer t such that for any a and b summing to t, H can be partitioned into a stable sets and b cliques.

Remark: $t(H)$ is clearly at least the chromatic number of H, at least the chromatic number of the complement of H, and at most the sum of these two numbers.

Keyest Lemma: For every $½>d>0$ and graph H, there are $e_{d,H}>0$, $N_{d,H}$, and $a_{d,H}>0$ such that if $t(H)$ is l, and $X_1,\ldots,X_l$ are subsets of $V(G)$, each of size $N \geq N_{d,H}$ and for all $1\leq i<j\leq l$, the pair (X_i,X_j) is $e_{d,H}$-regular with density which is at least d and at most 1-d then G contains at least $a_{d,H}N^{\wedge|V(H)|}$ induced copies of H.

Proof: We need an auxiliary result which we prove below:

Refinement Lemma: for every $k>1,1/2>d>0$, $e>0$ there is an $N_{k,e}$ and $S_{k,e}$ such that if G has at least $N_{k,e}$ vertices then there is an e-regular partition $(X_0,X_1,\ldots,X_p)$ of G where p is at at most $S_{k,e}$ such that for some k of these parts, either (i) every pair is e-regular with density at least d, or (ii) every pair is e-regular with density at most d.

Proof of Keyest Lemma: We consider $N_{d/2,|V(H)|}$,$e=e_{d/2,|V(H)|}$,$a_{d/2,|V(H)|}$ as in the Keyer Lemma. We set k to be $|V(H)|$ and consider $S_{k,e}$ as in the refinement lemma. We set $N_{d,H}$ to be $N_{d,|V(H)|}S_{k,e}/(1-e)$.
We set $e_{d,H}$ to be $(1-e)e/S_{k,e}$ We set $a_{d,H}$ to be $a_{d/2,|V(H)|}(S_{k,e}/1-e)^{|V(H)|}$.

We apply the Refinement Lemma to each X_i. WLOG, we can assume that for some $a<t(G)$, the partions of $X_1,\ldots,X_a$ satisfy (i), while the remaining partitions satisfy (ii). We consider a partition of G into cliques $S_1,\ldots,S_a$ and stable sets $S_{a+1},\ldots,S_l$. For each i, we choose $|S_i|$ of the k special parts of X_i whose existence is guaranteed by the refinement lemma and a bijection between these parts and the vertices of H in S_i. For each vertex v_i of H, we let Z_i be the corresponding part in the bijection. We can easily verify that the Z_i verify the hypotheses of the Keyer Lemma for $|V(H)|$ and d/2. So, applying that lemma we obtain the desired result. QED.

Proof of the Refinement Lemma: We can assume that e is at most 10^{-6k}, as if the result is true for e it is true for e'>e. We let k' be 10^{2k+1} and set $S_{k,e}$ to be the $M_{k',e}$ given by the Regularity Lemma. We apply the Regularity lemma with e and k'. From the resultant parts, we choose a set S_0 of p/10 which are in the fewest irregular pairs. It follows each is in at most $p/10^{3k}$ irregular pairs. We now proceed for a sequence of 2k iterations. After i iterations we will have construced a sequence $Y_i,\ldots Y_i$ of parts and a set S_i of at least $p/10^{i+1}$ parts such that:

For each j between 1 and i, we have either (i) for every part X which is in S_i or is $Y_j$for some j>i, (X_jY) is e-regular with density>d, or (ii) for every part X which is in S_i or is $Y_j$for some j>i, (X_jY) is e-regular with density<1-d.
To extend this sequence in iteration i+1, we consider any part X_{i+1} in S_i. It is irregular with at most $p/10^{3k}<|S_i|/10$ of the parts in S_i. Since each of the other pairs containing it either has density exceeding d or less than 1-d, we can clearly choose S_{i+1} in S_i so that (i) or (ii) holds for j=i+1 and hence for all j.
Now, we have k of the Y_i for which (i) holds, or k for which (ii) holds. QED.