ON COMPUTING THE FREQUENCIES OF INDUCED SUBHYPERGRAPHS

BRENDAN NAGLE

Abstract. Let $\mathcal{F}$ be an r-uniform hypergraph with f vertices, where $f > r \geq 3$. In [12], R. Yuster posed the problem of whether there exists an algorithm which, for a given r-uniform hypergraph $\mathcal{H}$ with n vertices, computes the number of induced copies of $\mathcal{F}$ in $\mathcal{H}$ in time $o(n^f)$. The analogous question for graphs ($r = 2$) was known to hold from a $O(n^{f-\varepsilon})$ time algorithm of Nešetřil and Poljak [9] (for a constant $\varepsilon = \varepsilon_f > 0$ which is independent of n). Here, we present an algorithm for this problem, when $r \geq 3$, with running time $O(n^f / \log_2 n)$.

1. Introduction

In this paper, we consider algorithms for computing the number of copies of a fixed r-uniform hypergraph $\mathcal{F}$ which are induced subhypergraphs of a given r-uniform hypergraph $\mathcal{H}$. Let $\mathcal{F}$ have f vertices and let $\mathcal{H}$ have vertex set $V = V(\mathcal{H})$. We write $\mathcal{F}_{\text{ind}}(\mathcal{H})$ for the collection of all f-element vertex subsets $S \in \binom{V}{f}$ which induce a copy of $\mathcal{F}$ in $\mathcal{H}$. (Note that $\binom{V}{f}$ denotes the family of all f-element subsets of V.) Elements of $\mathcal{F}_{\text{ind}}(\mathcal{H})$ correspond to *unlabeled* induced copies of $\mathcal{F}$ in $\mathcal{H}$. (We discuss labeled as well as not-necessarily induced copies below.) When $\mathcal{F} = K_f^{(r)}$ is the *f-clique*, the complete r-uniform hypergraph on f vertices, we write $K_f^{(r)}(\mathcal{H})$ for $\mathcal{F}_{\text{ind}}(\mathcal{H})$, and refer to $K_{r+1}^{(r)}$ as the *r-simplex*.

In the case of graphs ($r = 2$), Nesětřil and Poljak [9] gave an algorithm that uses fast matrix multiplication to determine $|\mathcal{F}_{\text{ind}}(\mathcal{H})|$ in time $O\big(n^{\omega\lfloor f/3\rfloor+(f \bmod 3)}\big)$, where $\omega \leq 2.376$ (see [2]) is the exponent of matrix multiplication. In the course of studying this and several related problems for hypergraphs, R. Yuster [12, 13] formulated the following problem (see Problem 6.1 of [12]).

Problem 1.1 (Yuster [12, 13]). *Let $\mathcal{F}$ be an r-uniform hypergraph with $f > r \geq 3$ vertices. Is there an algorithm which, for a given r-uniform hypergraph $\mathcal{H}$ with n vertices, computes $|\mathcal{F}_{\text{ind}}(\mathcal{H})|$ in time $o(n^f)$? In particular, when $\mathcal{F} = K_{r+1}^{(r)}$ is the r-simplex, is there an algorithm which, in time $o(n^{r+1})$, determines if $|K_{r+1}^{(r)}(\mathcal{H})| > 0$?*

In this paper, we present such an algorithm. (All logarithms in this paper are taken base 2.)

Theorem 1.2. *Let $\mathcal{F}$ be an r-uniform hypergraph with $f > r \geq 3$ vertices. There exists an algorithm $\boldsymbol{A}_{\mathcal{F}}$ which, for a given r-uniform hypergraph $\mathcal{H}$ with n vertices, computes the quantity $|\mathcal{F}_{\text{ind}}(\mathcal{H})|$ in time $O(n^f / \log n)$. Moreover, $\boldsymbol{A}_{\mathcal{F}}$ finds an induced copy of $\mathcal{F}$ in $\mathcal{H}$ whenever there is one.*

Theorem 1.2 admits, as corollaries, algorithms for counting the number of *labeled* copies of $\mathcal{F}$ in $\mathcal{H}$, in both the induced and not-necessarily induced cases (summarized below in Corollary 1.3). For the induced case, let $\vec{\mathcal{F}}_{\text{ind}}(\mathcal{H})$ denote the family of all injections $\psi : V(\mathcal{F}) \to V(\mathcal{H})$

Research partially supported by NSF grant DMS 0639839.

satisfying that, for each r-tuple $R \in \binom{V(\mathcal{F})}{r}$, $\psi(R) \in \mathcal{H}$ if, and only if, $R \in \mathcal{F}$. Note that $\vec{\mathcal{F}}_{\mathrm{ind}}(\mathcal{F}) = \mathrm{Aut}(\mathcal{F})$ corresponds to the automorphism group of $\mathcal{F}$, the size of which is computable in constant time. Since $|\vec{\mathcal{F}}_{\mathrm{ind}}(\mathcal{H})| = |\mathrm{Aut}(\mathcal{F})| \times |\mathcal{F}_{\mathrm{ind}}(\mathcal{H})|$, Theorem 1.2 implies $|\vec{\mathcal{F}}_{\mathrm{ind}}(\mathcal{H})|$ is computable in time $O(n^f/\log n)$. For the not-necessarily induced case, let $\vec{\mathcal{F}}(\mathcal{H})$ denote the family of all injections $\psi : V(\mathcal{F}) \to V(\mathcal{H})$ satisfying that $\psi(R) \in \mathcal{H}$ for each $R \in \mathcal{F}$. To compute $|\vec{\mathcal{F}}(\mathcal{H})|$, let $\boldsymbol{\mathcal{F}}$ denote the family of all 'superhypergraphs' $\mathcal{G} \supseteq \mathcal{F}$ on vertex set $V(\mathcal{F})$. For $\mathcal{G}_1, \mathcal{G}_2 \in \boldsymbol{\mathcal{F}}$, let $\mathcal{G}_1 \sim \mathcal{G}_2$ if, and only if, $\mathcal{G}_1$ and $\mathcal{G}_2$ are isomorphic, and let $\boldsymbol{\mathcal{F}}_\sim$ be a class of representatives from the partition of $\boldsymbol{\mathcal{F}}$ induced by the equivalence relation $\sim$ (which is constructable in constant time.) Then $|\vec{\mathcal{F}}(\mathcal{H})| = \sum_{\mathcal{G} \in \boldsymbol{\mathcal{F}}_\sim} |\vec{\mathcal{G}}_{\mathrm{ind}}(\mathcal{H})|$, where each of these terms is computable in time $O(n^f/\log n)$.

Corollary 1.3. *Let $\mathcal{F}$ be an r-uniform hypergraph with $f > r \geq 3$ vertices. There exist algorithms which, for a given r-uniform hypergraph $\mathcal{H}$ with n vertices, compute the quantities $|\vec{\mathcal{F}}_{\mathrm{ind}}(\mathcal{H})|$ and $|\vec{\mathcal{F}}(\mathcal{H})|$ in time $O(n^f/\log n)$.*

Algorithms for closely approximating $|\mathcal{F}_{\mathrm{ind}}(\mathcal{H})|$ can have significantly lower complexity than their exact counterparts. In the case of graphs ($r = 2$), Duke, Lefmann and Rödl [4] gave a $O(n^{2.376})$ algorithm for approximating $|\mathcal{F}_{\mathrm{ind}}(\mathcal{H})|$ within an error of $o(n^f)$. This algorithm is based on an algorithmic version of the celebrated Szemerédi regularity lemma [10, 11] given by Alon, Duke, Lefmann, Rödl and Yuster [1] (also considered in [4]). Kohayakawa, Rödl and Thoma [7] later improved the running time of [4] to $O(n^2)$ by establishing an improved constructive version of the regularity lemma. In the case of 3-uniform hypergraphs, Haxell, Nagle and Rödl [6] established a $O(n^6)$ algorithm approximating $|\mathcal{F}_{\mathrm{ind}}(\mathcal{H})|$ within an error of $o(n^f)$. This algorithm is based on an algorithmic version of a hypergraph regularity lemma of Frankl and Rödl [5] (cf. [3, 8]).

To conclude this introduction, we believe that it would be interesting to improve the exponent of computing $|\mathcal{F}_{\mathrm{ind}}(\mathcal{H})|$ for a fixed but arbitrary r-uniform hypergraph $\mathcal{F}$.

Problem 1.4. *For each r-uniform hypergraph $\mathcal{F}$ with $f > r \geq 3$ vertices, do there exist $\varepsilon = \varepsilon(\mathcal{F}) > 0$ and an algorithm $\hat{\boldsymbol{A}}_{\mathcal{F}}$ which, for a given r-uniform hypergraph $\mathcal{H}$ with n vertices, computes the quantity $|\mathcal{F}_{\mathrm{ind}}(\mathcal{H})|$ in time $O(n^{f-\varepsilon})$?*

Our paper is organized as follows. The heart of the proof of Theorem 1.2 concerns the special case when $\mathcal{F} = K_{r+1}^{(r)}$ is the r-simplex, to which we devote Section 2. Section 3 handles all remaining details of Theorem 1.2.

Acknowledgement. The author would like to thank the referees for their careful reading, and for their very helpful comments which improved our exposition.

2. Proof of Theorem 1.2 for r-simplices

For $r \geq 3$, let r-uniform hypergraph $\mathcal{H}$ be given on vertex set V, where $|V| = n$. We shall assume that $\mathcal{H}$ is represented by its characteristic function $\chi_{\mathcal{H}} : \binom{V}{r} \to \{0, 1\}$, where for a given $R \in \binom{V}{r}$, $\chi_{\mathcal{H}}(R) = 1$ if, and only if, $R \in \mathcal{H}$. We establish the algorithm $\boldsymbol{A}_r$ which computes $|K_{r+1}^{(r)}(\mathcal{H})|$ in time $O(n^{r+1}/\log n)$. At the end of the section, it will be easy to indicate how $\boldsymbol{A}_r$ can also find an r-simplex in $\mathcal{H}$, when there is one. We now describe the first (and main) step of the algorithm $\boldsymbol{A}_r$.

Step 1. Let $U_1 = \{u_1, \ldots, u_m\} \in \binom{V}{m}$ be an arbitrary set of m vertices, where $m = \lfloor (1/2) \log n \rfloor$. The main goal of Step 1 is to count the number of r-simplices in $\mathcal{H}$ having at least one vertex in U_1. To that end, let $K^{(r)}_{r+1}(U_1, \mathcal{H})$ denote the collection of sets $S \in \binom{V}{r+1}$ which span an r-simplex in $\mathcal{H}$ and which satisfy $S \cap U_1 \neq \emptyset$. We assert the following.

Proposition 2.1. *The quantity $|K^{(r)}_{r+1}(U_1, \mathcal{H})|$ can be computed in time $O(n^r)$.*

Theorem 1.2 (for r-simplices) now follows by iterating Step 1. Indeed, let $V_1 = V \setminus U_1$ and $\mathcal{H}_1 = \mathcal{H}[V_1]$, where $\mathcal{H}[V_1]$ is the subhypergraph of $\mathcal{H}$ induced on $V_1 = V \setminus U_1$. Let $U_2 \subseteq V_1$ be a subset of size m. Step 2 computes, in time $O(n^r)$, the number $|K^{(r)}_{r+1}(U_2, \mathcal{H}_1)|$ of r-simplices in $\mathcal{H}_1$ having at least one vertex in U_2. (Note that $K^{(r)}_{r+1}(U_2, \mathcal{H}_1)$ and $K^{(r)}_{r+1}(U_1, \mathcal{H})$ are disjoint.) Repeating[1] this procedure n/m times computes all of $|K^{(r)}_{r+1}(\mathcal{H})|$, in time $O(n^{r+1}/m) = O(n^{r+1}/\log n)$, as promised.

Proof of Proposition 2.1. We first perform a greedy process, so that the remainder of the proof addresses only essential details. Note that the elements $S \in K^{(r)}_{r+1}(U_1, \mathcal{H})$ fall into two classes: those for which $|S \cap U_1| \geq 2$, and those for which $|S \cap U_1| = 1$. Let $\#_{U_1}$ denote the size of the former class, which can be greedily computed in time $O\left(m^2 n^{r-1}\right) = O(n^{r-1} \log^2 n) = o(n^r)$. We now determine $|K^{(r)}_{r+1}(U_1, \mathcal{H})| - \#_{U_1}$, which counts the elements $S \in K^{(r)}_{r+1}(U_1, \mathcal{H})$ meeting U_1 exactly once. We begin with a sketch of the approach.

Sketch. Observe that

$$|K^{(r)}_{r+1}(U_1, \mathcal{H})| - \#_{U_1} = \sum \{\deg_{U_1}(H) : H \in \mathcal{H}_1\}\,, \tag{1}$$

where for $H \in \mathcal{H}_1$, $\deg_{U_1}(H) = |\{u \in U_1 : \{u\} \cup H \in K^{(r)}_{r+1}(U_1, \mathcal{H})\}|$. Our plan is to construct, in time $O(n^r)$, a partition $\Pi_{\mathcal{H}_1}$ of $\mathcal{H}_1$ into $O(n^{r-1})$ classes $\mathbf{H} \in \Pi_{\mathcal{H}_1}$ with the property that, whenever $H, H' \in \mathbf{H} \in \Pi_{\mathcal{H}_1}$, then $\deg_{U_1}(H) = \deg_{U_1}(H')$. In this way, for each class $\mathbf{H} \in \Pi_{\mathcal{H}_1}$, we have that $\deg_{U_1}(\mathbf{H})$ is constant, and so computed in time $O(m)$. Therefore, the degrees $\deg_{U_1}(\mathbf{H})$, over all $O(n^{r-1})$ classes $\mathbf{H} \in \Pi_{\mathcal{H}_1}$, are computed in time $O(mn^{r-1}) = O(n^{r-1} \log n) = o(n^r)$. We then compute (1) by

$$|K^{(r)}_{r+1}(U_1, \mathcal{H})| - \#_{U_1} = \sum \{\deg_{U_1}(\mathbf{H}) \times |\mathbf{H}| : \mathbf{H} \in \Pi_{\mathcal{H}_1}\}\,. \tag{2}$$

Note that we may assume the sizes $|\mathbf{H}|$ are computed when $\Pi_{\mathcal{H}_1}$ was constructed. (Alternatively, once $\Pi_{\mathcal{H}_1}$ is constructed, we may construct the list $\{|\mathbf{H}| : \mathbf{H} \in \Pi_{\mathcal{H}_1}\}$ of sizes in time $\sum \{O(|\mathbf{H}|) : \mathbf{H} \in \Pi_{\mathcal{H}_1}\} = O(|\mathcal{H}_1|) = O(n^r)$.) This completes the sketch of the proof. □

What essentially remains is to construct the partition $\Pi_{\mathcal{H}_1}$, for which we now prepare. To that end, we first construct the following partition $\Pi^{(r-1)}_{V_1}$ of $\binom{V_1}{r-1}$, the family of $(r-1)$-tuples $Q \in \binom{V_1}{r-1}$ from $V_1 = V \setminus U_1$. To describe $\Pi^{(r-1)}_{V_1}$, consider the mapping $\psi : \binom{V_1}{r-1} \to \{0,1\}^m$ defined by, for $Q \in \binom{V_1}{r-1}$,

$$\psi(Q) = \boldsymbol{u}_Q = \left(\chi_{\mathcal{H}}(\{u_1\} \cup Q), \ldots, \chi_{\mathcal{H}}(\{u_m\} \cup Q)\right) \in \{0,1\}^m\,, \tag{3}$$

[1]More formally, repeat this procedure almost n/m times, until say $O(\sqrt{n})$ vertices remain, and finish the job by exhaustively searching for the remaining r-simplices.

where recall that $\chi_{\mathcal{H}}$ is the characteristic function for $\mathcal{H}$ and that $U_1 = \{u_1, \ldots, u_m\}$. Now, for $\boldsymbol{u} \in \{0,1\}^m$, set

$$\mathcal{Q}_{\boldsymbol{u}} = \psi^{-1}(\boldsymbol{u}) = \left\{ Q \in \tbinom{V_1}{r-1} : \boldsymbol{u}_Q = \boldsymbol{u} \right\},$$

which is an $(r-1)$-uniform hypergraph on vertex set V_1. The promised partition $\Pi_{V_1}^{(r-1)}$ of $\binom{V_1}{r-1}$ is then

$$\Pi_{V_1}^{(r-1)} = \left\{ \mathcal{Q}_{\boldsymbol{u}} : \boldsymbol{u} \in \{0,1\}^m \right\}.$$

We claim that ψ and $\Pi_{V_1}^{(r-1)}$ may be constructed in time $O(n^{r-1}\log n)$. Indeed, first construct the space $\{0,1\}^m$ in time $O(2^m) = O(\sqrt{n}) = o(n^{r-1})$ (recall $2m \le \log n$ and $r \ge 3$). Then, construct ψ in time $O(n^{r-1}m) = O(n^{r-1}\log n)$. Observe that $\Pi_{V_1}^{(r-1)}$ is constructed from ψ in time $\sum \left\{ O(|\psi^{-1}(\boldsymbol{u})|) : \boldsymbol{u} \in \{0,1\}^m \right\} = O(n^{r-1})$ (note that at most $O(\sqrt{n}) = o(n^{r-1})$ zero terms are considered).

To construct the promised partition $\Pi_{\mathcal{H}_1}$, we now consider the following mapping $\varrho : \mathcal{H}_1 \to \prod_{k=1}^r \{0,1\}^m = \{0,1\}^m \times \cdots \times \{0,1\}^m$. For $H = \{v_1, \ldots, v_r\} \in \mathcal{H}_1$, write $\binom{H}{r-1} = \{Q_1, \ldots, Q_r\}$, where $Q_i = H \setminus \{v_i\}$, for all $1 \le i \le r$. Define

$$\varrho(H) = (\boldsymbol{u}_{Q_1}, \ldots, \boldsymbol{u}_{Q_r}) = (\psi(Q_1), \ldots, \psi(Q_r)) \in \prod_{k=1}^r \{0,1\}^m,$$

where ψ is the mapping constructed above for the partition $\Pi_{V_1}^{(r-1)}$. Now, for $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r) \in \prod_{k=1}^r \{0,1\}^m$, let

$$\mathcal{H}_{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_r} = \varrho^{-1}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r) = \left\{ H \in \mathcal{H}_1 : \varrho(H) = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r) \right\}.$$

(The object $\mathcal{H}_{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_r}$, an r-uniform subhypergraph of $\mathcal{H}_1$, is a class $\mathbf{H}$ from the Sketch.) The promised partition $\Pi_{\mathcal{H}_1}$ (from the Sketch) is then

$$\Pi_{\mathcal{H}_1} = \left\{ \mathcal{H}_{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_r} : (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r) \in \prod_{k=1}^r \{0,1\}^m \right\}.$$

Note that $\Pi_{\mathcal{H}_1}$ is a partition with at most

$$\left|\Pi_{\mathcal{H}_1}\right| \le 2^{mr} \le n^{r/2} = O(n^{r-1}) \tag{4}$$

classes, since $2m \le \log n$ and $r \ge 3$. We claim that both ϱ and $\Pi_{\mathcal{H}_1}$ may be constructed in time $O(n^r)$. Indeed, to construct ϱ, first construct the space $\prod_{k=1}^r \{0,1\}^m$ in time $O(2^{mr}) = O(n^{r-1})$. Then, for each $H \in \mathcal{H}_1$, where $\binom{H}{r-1} = \{Q_1, \ldots, Q_r\}$, one recalls $(\psi(Q_1), \ldots, \psi(Q_r))$ in constant time (cf. (3)). Thus, ϱ is constructed in time $O(n^r)$. From ϱ, one constructs $\Pi_{\mathcal{H}_1}$ in time

$$\sum \left\{ O(|\varrho^{-1}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r)|) : (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r) \in \prod_{k=1}^r \{0,1\}^m \right\} = O(|\mathcal{H}_1|) = O(n^r) \tag{5}$$

(note that at most $O(n^{r-1})$ zero terms are considered). For future reference, let us also now compute,

$$\text{for each } (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r) \in \textstyle\prod_{k=1}^r \{0,1\}^m, \text{ the size } |\mathcal{H}_{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_r}|, \tag{6}$$

which can be done simultaneously in (5).

We return to (1) and (2), and consider $|K_{r+1}^{(r)}(U_1, \mathcal{H})| - \#_{U_1} = \sum_{H \in \mathcal{H}_1} \deg_{U_1}(H)$. The following claim addresses how we compute degrees in this sum.

Claim 2.2. *Fix $H \in \mathcal{H}_{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_r} \in \Pi_{\mathcal{H}_1}$.*

(1) *If, for some* $1 \le i \le m$*, the projection* π_i *onto the* i^{th} *coordinate satisfies* $\prod_{j=1}^{r} \pi_i(\boldsymbol{u}_j) = 1$*, then* $\{u_i\} \cup H \in K_{r+1}^{(r)}(\mathcal{H})$*;*
(2) $\deg_{U_1}(H) = \sum_{i=1}^{m} \prod_{j=1}^{r} \pi_i(\boldsymbol{u}_j)$*;*
(3) $\deg_{U_1}(H)$ *does not depend on* $H \in \mathcal{H}_{\boldsymbol{u}_1,\dots,\boldsymbol{u}_r}$*, but only on the class* $\mathcal{H}_{\boldsymbol{u}_1,\dots,\boldsymbol{u}_r}$ *to which* H *belongs.*

Proof. Let $H = \{v_1, \dots, v_r\} \in \mathcal{H}_{\boldsymbol{u}_1,\dots,\boldsymbol{u}_r} \in \Pi_{\mathcal{H}_1}$ be given, where we assume that for each $1 \le j \le r$, where $Q_j = H \setminus \{v_j\}$, we have $\boldsymbol{u}_{Q_j} = \boldsymbol{u}_j$. Now, suppose some $1 \le i \le m$ satisfies $\prod_{j=1}^{r} \pi_i(\boldsymbol{u}_j) = 1$. Then, for each $1 \le j \le r$, we have $\pi_i(\boldsymbol{u}_j) = 1 = \chi(\{u_i\} \cup Q_j)$ (recall χ is the characteristic function of $\mathcal{H}$), in which case $\{u_i\} \cup Q_j \in \mathcal{H}$. Since this holds for every $1 \le j \le r$, then together with H, $\{u_i\} \cup H$ spans an r-simplex $K_{r+1}^{(r)}$ in $\mathcal{H}$. The second assertion now follows from the first, and third assertion follows from the second. ☐

We conclude the proof of Proposition 2.1. For a class $\mathcal{H}_{\boldsymbol{u}_1,\dots,\boldsymbol{u}_r} \in \Pi_{\mathcal{H}_1}$, the quantity $\sum_{i=1}^{m} \prod_{j=1}^{r} \pi_i(\boldsymbol{u}_j)$ is a multilinear form (generalizing the dot product), which we abbreviate to $\langle \boldsymbol{u}_1, \dots, \boldsymbol{u}_r \rangle$. Then, from (2) and Claim 2.2, we see that $\deg_{U_1}(\mathcal{H}_{\boldsymbol{u}_1,\dots,\boldsymbol{u}_r}) = \langle \boldsymbol{u}_1, \dots, \boldsymbol{u}_r \rangle$. As such, from (1), (2) and Claim 2.2,

$$|K_{r+1}^{(r)}(U_1, \mathcal{H})| - \#_{U_1} = \sum_{H \in \mathcal{H}_1} \deg_{U_1}(H) = \sum \left\{ \langle \boldsymbol{u}_1, \dots, \boldsymbol{u}_r \rangle \times |\mathcal{H}_{\boldsymbol{u}_1,\dots,\boldsymbol{u}_r}| : \mathcal{H}_{\boldsymbol{u}_1,\dots,\boldsymbol{u}_r} \in \Pi_{\mathcal{H}_1} \right\}. \tag{7}$$

Since $\Pi_{\mathcal{H}_1}$ was already constructed (recall (6)), the sum in (7) is computed in an additional time of $O(mn^{r-1}) = O(n^{r-1} \log n)$. Indeed, for each $\mathcal{H}_{\boldsymbol{u}_1,\dots,\boldsymbol{u}_r} \in \Pi_{\mathcal{H}_1}$, $\langle \boldsymbol{u}_1, \dots, \boldsymbol{u}_r \rangle$ requires $O(m)$ computations, and $\Pi_{\mathcal{H}_1}$ consists of $O(n^{r-1})$ elements (recall (4)). This completes the proof of Proposition 2.1.

On finding an r-simplex in $\mathcal{H}$. The algorithm $\boldsymbol{A}_r$ will find an r-simplex in $\mathcal{H}$ when there is one. Indeed, suppose that Step i (recall Step 1), $1 \le i = O(n/\log n)$, is the first for which $\boldsymbol{A}_r$ determines that $|K_{r+1}^{(r)}(\mathcal{H})| > 0$. Without loss of generality, suppose $i = 1$. Recall that $\boldsymbol{A}_r$ performs an exhaustive search to compute $\#_{U_1}$ (recall (1)), and so it could return the first instance it finds verifying that $\#_{U_1} > 0$. Suppose, otherwise, that $\#_{U_1} = 0$ so that, for some (first) class $\mathcal{H}_{\boldsymbol{u}_1,\dots,\boldsymbol{u}_r} \in \Pi_{\mathcal{H}_1}$ (cf. (7)), the algorithm determines that both $\langle \boldsymbol{u}_1, \dots, \boldsymbol{u}_r \rangle, |\mathcal{H}_{\boldsymbol{u}_1,\dots,\boldsymbol{u}_r}| > 0$. Then, let $1 \le i \le m$ be the first coordinate for which $\prod_{j=1}^{r} \pi_i(\boldsymbol{u}_j) = 1$. Then $\boldsymbol{A}_r$ takes any $H \in \mathcal{H}_{\boldsymbol{u}_1,\dots,\boldsymbol{u}_r}$ and returns $\{u_i\} \cup H$.

3. Proof of Theorem 1.2

The work that remains is quite standard, but we will consider separately the two cases when $\mathcal{F} = K_f^{(r)}$ is complete (an f-clique) and when $\mathcal{F}$ is not necessarily complete. In particular, we will first show how the algorithm $\boldsymbol{A}_r$ of the previous section can be extended to provide an algorithm $\boldsymbol{A}_{r,f}$ which computes, for a given r-uniform hypergraph $\mathcal{H}$ on n vertices, the quantity $|K_f^{(r)}(\mathcal{H})|$ in time $O(n^f/\log n)$. The algorithm $\boldsymbol{A}_{r,f}$ can also find an f-clique in $\mathcal{H}$ when there is one. Afterward, we will show, for an arbitrary r-uniform hypergraph $\mathcal{F}$ on f vertices, how the algorithm $\boldsymbol{A}_{r,f}$ can be extended to provide the promised algorithm $\boldsymbol{A}_{\mathcal{F}}$.

Algorithm $\boldsymbol{A}_{r,f}$. Fix an integer $r \ge 3$. We proceed by induction on $f \ge r+1$. When $f = r+1$, we take $\boldsymbol{A}_{r,r+1} = \boldsymbol{A}_r$ as the algorithm of the previous section. Now, for $f - 1 \ge r + 1$, assume there exists an algorithm $\boldsymbol{A}_{r,f-1}$ which, for a given r-uniform hypergraph $\mathcal{H}$ on n vertices, computes $|K_{f-1}^{(r)}(\mathcal{H})|$ in time $O(n^{f-1}/\log n)$. Suppose, moreover, that $\boldsymbol{A}_{r,f-1}$ can also find an $(f-1)$-clique in $\mathcal{H}$ when there is one. We now describe the promised algorithm $\boldsymbol{A}_{r,f}$.

Let $\mathcal{H}$ be a given r-uniform hypergraph with an n-element vertex set V, where we assume that $\mathcal{H}$ is represented by its characteristic function $\chi_{\mathcal{H}} : \binom{V}{r} \to \{0,1\}$. Fix an arbitrary vertex $u = u_1 \in V$, and construct the following two hypergraphs:

$$\mathcal{Q}_u = \left\{Q \in \tbinom{V \setminus \{u\}}{r-1} : \chi_{\mathcal{H}}(\{u_1\} \cup Q) = 1\right\}; \qquad \mathcal{H}_u = \mathcal{H} \cap K_r^{(r-1)}(\mathcal{Q}_u)\,.$$

Note that $\mathcal{Q}_u$ is an $(r-1)$-uniform hypergraph whose edges $Q \in \mathcal{Q}_u$, together with u, form an edge $H \in \mathcal{H}$. Note that $\mathcal{H}_u$ is an r-uniform hypergraph whose edges $H \in \mathcal{H}_u$ span an $(r-1)$-simplex $K_r^{(r-1)}$ in $\mathcal{Q}_u$. Clearly, $\mathcal{Q}_u$ can be constructed from $\chi_{\mathcal{H}}$ in time $O(n^{r-1})$. Note that $\mathcal{H}_u$ can be constructed from $\chi_{\mathcal{H}}$ in time $O(|\mathcal{H}|) = O(n^r)$. Indeed, for a fixed $H \in \mathcal{H}$, one computes $\chi_{\mathcal{H}}(\{u\} \cup Q)$ for each $Q \in \binom{H}{r-1}$.

Now, observe that the quantity $|K_{f-1}^{(r)}(\mathcal{H}_u)|$ counts the number of cliques $K_f^{(r)}$ in $\mathcal{H}$ which contain the vertex u. By induction, $\boldsymbol{A}_{r,f-1}$ counts $K_{f-1}^{(r)}(\mathcal{H}_u)$ in time $O(n^{f-1}/\log n)$. Moreover, $\boldsymbol{A}_{r,f-1}$ finds an $(f-1)$-clique in $\mathcal{H}_u$, if there is one, which combined with u forms an f-clique in $\mathcal{H}$. We repeat this procedure for a vertex $u_2 \in V \setminus \{u\}$ for the hypergraph $\mathcal{H}[V \setminus \{u\}]$, and so on. After n iterations, we have counted all of $K_f^{(r)}(\mathcal{H})$ in time $O(n^f/\log n)$, and have found an f-clique in $\mathcal{H}$, if there is one. This describes the algorithm $\boldsymbol{A}_{r,f}$.

Algorithm $\boldsymbol{A}_{\mathcal{F}}$. Let r-uniform hypergraph $\mathcal{F}$ on $f > r \geq 3$ vertices be given. Let $\mathcal{H}$ be a given r-uniform hypergraph with an n-element vertex set V. (In this proof, we make only tacit use of the characteristic function representing $\mathcal{H}$.) We begin with a greedy procedure, so that only essential details remain. For $t = \lfloor \log n \rfloor$, construct (in linear time) any partition $V(\mathcal{H}) = V_1 \cup \cdots \cup V_t$ for which $|V_1| \leq \cdots \leq |V_t| \leq |V_1| + 1$. Employ an exhaustive search for all elements $S \in \mathcal{F}_{\text{ind}}(\mathcal{H})$ for which $|S \cap V_i| \geq 2$ for some $1 \leq i \leq t$, which clearly may be completed in time $O(t\binom{\lceil n/t \rceil}{2} n^{f-2}) = O(n^f/\log n)$. If any such S is found in this search, return the first one for the promised example of an induced copy of $\mathcal{F}$ in $\mathcal{H}$.

The remainder of $\boldsymbol{A}_{\mathcal{F}}$ will count *crossing* copies $S \in \mathcal{F}_{\text{ind}}(\mathcal{H})$, that is, f-tuples $S \in \mathcal{F}_{\text{ind}}(\mathcal{H})$ for which $|S \cap V_i| \leq 1$ for all $1 \leq i \leq t$. Write $\mathcal{F}_{\text{ind}}^{\times}(\mathcal{H})$ for the family of all crossing copies $S \in \mathcal{F}_{\text{ind}}(\mathcal{H})$. In what follows, we shall consider two partitions of $\mathcal{F}_{\text{ind}}(\mathcal{H})$ (see upcoming (8) and (9)), where the second partition (9) will refine the first (8). These partitions provide us an identity in upcoming (10) for computing $|\mathcal{F}_{\text{ind}}^{\times}(\mathcal{H})|$. Our next several efforts will be to describe these partitions. At the end of the section, we handle all remaining constructive details of $\boldsymbol{A}_{\mathcal{F}}$.

First partition. For $F = \{i_1, \ldots, i_f\} \in \binom{[t]}{f}$, consider

$$\mathcal{F}_{\text{ind}}^{\times}(\mathcal{H}; F) = \mathcal{F}_{\text{ind}}^{\times}(\mathcal{H}) \cap \mathcal{F}_{\text{ind}}(\mathcal{H}[V_{i_1}, \ldots, V_{i_f}])\,,$$

where $\mathcal{H}[V_{i_1}, \ldots, V_{i_f}]$ is the f-partite subhypergraph of $\mathcal{H}$ induced by the partition $V_{i_1} \cup \cdots \cup V_{i_f}$. In other words, $\mathcal{F}_{\text{ind}}^{\times}(\mathcal{H}; F)$ is the family of all crossing $S \in \mathcal{F}_{\text{ind}}^{\times}(\mathcal{H})$ whose vertices lie within $V_{i_1} \cup \cdots \cup V_{i_f}$. Then

$$\mathcal{F}_{\text{ind}}^{\times}(\mathcal{H}) = \bigcup \left\{\mathcal{F}_{\text{ind}}^{\times}(\mathcal{H}; F) : F \in \tbinom{[t]}{f}\right\} \tag{8}$$

is a partition. We shall consider a refinement of (8) (in upcoming (9)), but will first require a few preparations.

Fix $F_0 = \{i_1, \ldots, i_f\} \in \binom{[t]}{f}$, and define $\mathbb{F}_{F_0}$ to be the family of $f!/|\text{Aut}(\mathcal{F})|$ many distinct (unlabeled) copies of $\mathcal{F}$ on vertex set F_0. Fix a copy $\mathcal{F}_0 \in \mathbb{F}_{F_0}$, and consider the following

r-partite r-uniform hypergraph $\mathcal{G} = \mathcal{G}_{F_0,\mathcal{F}_0}$ with vertex partition $V_{i_1} \cup \cdots \cup V_{i_f}$. For each $R = \{j_1, \ldots, j_r\} \in \binom{F_0}{r}$, define

$$\mathcal{G}_R = \mathcal{G}_{F_0,\mathcal{F}_0,R} = \begin{cases} \mathcal{H}[V_{j_1}, \ldots, V_{j_r}] & \text{if } R \in \mathcal{F}_0, \\ K^{(r)}(V_{j_1}, \ldots, V_{j_r}) \setminus \mathcal{H} & \text{if } R \notin \mathcal{F}_0, \end{cases}$$

where $K^{(r)}(V_{j_1}, \ldots, V_{j_r})$ is the complete r-partite r-uniform hypergraph with vertex partition $V_{j_1} \cup \cdots \cup V_{j_r}$. Define

$$\mathcal{G} = \mathcal{G}_{F_0,\mathcal{F}_0} = \bigcup \left\{ \mathcal{G}_R : \ R \in \binom{F_0}{r} \right\}.$$

We consider two properties of the hypergraph $\mathcal{G} = \mathcal{G}_{F_0,\mathcal{F}_0}$.

Fact 3.1. *Each element of $K_f^{(r)}(\mathcal{G}) = K_f^{(r)}(\mathcal{G}_{F_0,\mathcal{F}_0})$ corresponds to an element of $\mathcal{F}^{\times}_{\mathrm{ind}}(\mathcal{H}; F_0)$.*

Proof. Indeed, let $S = \{v_{i_1}, \ldots, v_{i_f}\} \in K_f^{(r)}(\mathcal{G})$, where $v_{i_1} \in V_{i_1}, \ldots, v_{i_f} \in V_{i_f}$. We claim that $i_j \mapsto v_{i_j}$, $1 \le j \le f$, defines an isomorphism from $\mathcal{F}_0$ to $\mathcal{H}[S] = \mathcal{H} \cap \binom{S}{r}$. Indeed, fix $R = \{j_1, \ldots, j_r\} \in \binom{F_0}{r}$. Suppose $R \in \mathcal{F}_0$. Then $\mathcal{G}_R = \mathcal{H}[V_{j_1}, \ldots, V_{j_r}]$, and so $\{v_{j_1}, \ldots, v_{j_r}\} \in \mathcal{G}_R$ implies $\{v_{j_1}, \ldots, v_{j_r}\} \in \mathcal{H}[S]$. Suppose $R \notin \mathcal{F}_0$. Then, $\mathcal{G}_R = K^{(r)}(V_{j_1}, \ldots, V_{j_r}) \setminus \mathcal{H}$, and so $\{v_{j_1}, \ldots, v_{j_r}\} \in \mathcal{G}_R$ implies $\{v_{j_1}, \ldots, v_{j_r}\} \notin \mathcal{H}[S]$. □

Fact 3.2. *For every element $S \in \mathcal{F}^{\times}_{\mathrm{ind}}(\mathcal{H}; F_0)$, there exists a unique $\mathcal{F}_S \in \mathbb{F}_{F_0}$ for which $S \in K_f^{(r)}(\mathcal{G}_{F_0,\mathcal{F}_S})$.*

Proof. Indeed, let $S = \{v_{i_1}, \ldots, v_{i_f}\} \in \mathcal{F}^{\times}_{\mathrm{ind}}(\mathcal{H}; F_0)$ be given, where $v_{i_1} \in V_{i_1}, \ldots, v_{i_f} \in V_{i_f}$. Let $v_{i_j} \mapsto i_j$, $1 \le j \le f$, be the isomorphism from $\mathcal{H}[S]$ (an induced copy of $\mathcal{F}$) to an element $\mathcal{F}_S \in \mathbb{F}_{F_0}$. We check that $S \in K_f^{(r)}(\mathcal{G}_{F_0,\mathcal{F}_S})$. Indeed, suppose $H = \{v_{j_1}, \ldots, v_{j_r}\} \in \mathcal{H}[S]$ so that $R = \{j_1, \ldots, j_r\} \in \mathcal{F}_S$. Then, $H \in \mathcal{H}[V_{j_1}, \ldots, V_{j_r}] = \mathcal{G}_{F_0,\mathcal{F}_S,R}$, and so $H \in \mathcal{G}_{F_0,\mathcal{F}_S}$. Suppose, on the other hand, that $H = \{v_{j_1}, \ldots, v_{j_r}\} \notin \mathcal{H}[S]$, so that $R = \{j_1, \ldots, j_r\} \notin \mathcal{F}_S$. Then $H \in K^{(r)}(V_{j_1}, \ldots, V_{j_r}) \setminus \mathcal{H} = \mathcal{G}_{F_0,\mathcal{F}_S,R}$, and so $H \in \mathcal{G}_{F_0,\mathcal{F}_S}$. The uniqueness assertion follows easily. Indeed, let $\mathcal{F}_1, \mathcal{F}_2 \in \mathbb{F}_{F_0}$ be given, where $R \in \mathcal{F}_1 \setminus \mathcal{F}_2$. Then, $\mathcal{G}_{F_0,\mathcal{F}_1,R}$ and $\mathcal{G}_{F_0,\mathcal{F}_2,R}$ are disjoint, in which case $K_f^{(r)}(\mathcal{G}_{F_0,\mathcal{F}_1})$ and $K_f^{(r)}(\mathcal{G}_{F_0,\mathcal{F}_2})$ are also disjoint. □

Second partition. It follows from the facts above that

$$\mathcal{F}^{\times}_{\mathrm{ind}}(\mathcal{H}; F_0) = \bigcup \left\{ K_f^{(r)}(\mathcal{G}_{F_0,\mathcal{F}_0}) : \ \mathcal{F}_0 \in \mathbb{F}_{F_0} \right\}$$

is a partition. Then

$$\mathcal{F}^{\times}_{\mathrm{ind}}(\mathcal{H}) = \bigcup \left\{ K_f^{(r)}(\mathcal{G}_{F,\mathcal{F}_0}) : \ F \in \binom{[t]}{f}, \ \mathcal{F}_0 \in \mathbb{F}_F \right\} \tag{9}$$

is a partition which refines (8). As such,

$$\left|\mathcal{F}^{\times}_{\mathrm{ind}}(\mathcal{H})\right| = \sum \left\{ \left|K_f^{(r)}(\mathcal{G}_{F,\mathcal{F}_0})\right| : \ F \in \binom{[t]}{f}, \ \mathcal{F}_0 \in \mathbb{F}_F \right\}. \tag{10}$$

To employ the identity in (10), we construct every hypergraph $\mathcal{G}_{F,\mathcal{F}_0}$, over all $F \in \binom{[t]}{f}$ and $\mathcal{F}_0 \in \mathbb{F}_F$. (Note that $\mathbb{F}_F$ is constructable in constant, viz. $O(f!)$, time.) For a fixed $F \in \binom{[t]}{f}$ and $\mathcal{F}_0 \in \mathbb{F}_F$, $\mathcal{G}_{F,\mathcal{F}_0}$ is constructed in time $O\big(\binom{f}{r}\lceil n/t \rceil^r\big)$. Therefore, all such hypergraphs $\mathcal{G}_{F,\mathcal{F}_0}$ are constructed in time

$$O\left(\binom{t}{f} f! \binom{f}{r} \lceil n/t \rceil^r\right) = O(n^r t^{f-r}) = O(n^r \log^{f-r} n) = o(n^f / \log n)$$

(recall $t = \lfloor \log n \rfloor$ and $f \geq r+1$). Now, we apply the algorithm $\boldsymbol{A}_{r,f}$ to each term in the sum in (10). Note that, for each $F \in \binom{[t]}{f}$ and $\mathcal{F}_0 \in \mathbb{F}_F$, the hypergraph $\mathcal{G}_{F,\mathcal{F}_0}$ has at most $f\lceil n/t \rceil$ vertices. Therefore, $\boldsymbol{A}_{r,f}$ computes $|K_f^{(r)}(\mathcal{G}_{F,\mathcal{F}_0})|$ in time $O\big((fn/t)^f / \log(fn/t)\big) = O(n^f/t^{f+1})$ (recall $t = \lfloor \log n \rfloor$). The time spent computing the terms $|K_f^{(r)}(\mathcal{G}_{F,\mathcal{F}_0})|$, over all $F \in \binom{[t]}{f}$ and $\mathcal{F}_0 \in \mathbb{F}_F$, is therefore

$$O\Big(\binom{t}{f} f! \frac{n^f}{t^{f+1}}\Big) = O(n^f/t) = O(n^f/\log n)\,.$$

This concludes our count of $|\mathcal{F}^{\times}_{\mathrm{ind}}(\mathcal{H})|$.

Note that we will find a copy $S \in \mathcal{F}^{\times}_{\mathrm{ind}}(\mathcal{H})$, if there is one, since $\boldsymbol{A}_{r,f}$ will find an f-clique in some $\mathcal{G}_{F,\mathcal{F}_0}$, for $F \in \binom{[t]}{f}$ and $\mathcal{F}_0 \in \mathbb{F}_F$. This completes our description of the algorithm $\boldsymbol{A}_{\mathcal{F}}$.

References

[1] Alon, N., Duke, R., Lefmann, H., Rödl, V., Yuster, R., *The algorithmic aspects of the Regularity Lemma*, J. Algorithms **16** (1994), 80–109.

[2] Coppersmith, D., Winograd, S., *Matrix multiplication via arithmetic progressions*, J. Symbolic Comput. **9** (1990), no. 3, 251–280.

[3] Dementieva, Y., Haxell, P., Nagle, B., Rödl, V., *On characterizing hypergraph regularity*, Random Structures Algorithms **21** (2002), no. 3–4, 293–332.

[4] Duke, R., Lefmann, H., Rödl, V., *A fast algorithm for computing the frequencies of subgraphs in a given graph*, SIAM J. Comp. **24** (1995), 598–620.

[5] Frankl, P., Rödl, V., *Extremal problems on set systems*, Random Structures Algorithms **20** (2002), no. 2, 131–164.

[6] Haxell, P.E., Nagle, B., Rödl, V., *An algorithmic version of the hypergraph regularity method*, SIAM J. Comput. **37** (2008), no. 6, 1728–1776.

[7] Kohayakawa, Y., Rödl, V., Thoma, L., *An optimal algorithm for checking regularity*, SIAM J. on Computing **32** (2003), no. 5, 1210–1235.

[8] Nagle, B., Poerschke, A., Rödl, V., Schacht, M., *Hypergraph regularity and quasirandomness*, in: Clair Mathieu (editor): Proceedings of the 20$^{\text{thm}}$ Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 09), pp. 227–245. ACM Press.

[9] Nešetřil, J., Poljak, S., *On the complexity of the subgraph problem*, Comment. Math. Univ. Carolin. **26** (1985), no. 2, 415–419.

[10] Szemerédi, E., *On sets of integers containing no k elements in arithmetic progression*, Acta Arithmetica **27** (1975), 199–245, Collection of articles in memory of Juriĭ Vladimirovič Linnik.

[11] Szemerédi, E., *Regular partitions of graphs*, Problèmes en combinatoires et théorie des graphes, (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS, vol. 260, CNRS, Paris, 1978, pp. 399–401.

[12] Yuster, R., *Finding and counting cliques and independent sets in r-uniform hypergraphs*, Information Processing Letters **99** (2006), 130–134.

[13] http://research.haifa.ac.il/∼raphy/problems.htm

Department of Mathematics and Statistics, University of South Florida, Tampa, FL 33620, USA.

E-mail address: bnagle@math.usf.edu.