The Algorithmic Aspects of Uncrowded Hypergraphs*†

Claudia Bertram-Kretzberg and Hanno Lefmann
Lehrstuhl Informatik II, Universität Dortmund, D-44221 Dortmund, Germany
bertram, lefmann@Ls2.informatik.uni-dortmund.de

Abstract

We consider the problem of finding deterministically a large independent set of guaranteed size in a hypergraph on n vertices and with m edges. With respect to the Turán bound, the quality of our solutions is for hypergraphs with not too many small cycles by a logarithmic factor in the input size better. The algorithms are fast; they often have a running time of $O(m) + o(n^3)$. Indeed, the denser the hypergraphs are the more close are the running times to the linear ones. This gives for the first time for some combinatorial problems algorithmic solutions with state-of-the-art quality, solutions of which so far only the existence was known. In some cases, the corresponding upper bounds match the lower bounds up to constant factors. The involved concepts are uncrowded hypergraphs.

1 Introduction

A fundamental problem in Computer Science and Mathematics is to find a large independent set in an arbitrary graph [6], [14], [20]. Recall, that for a graph $G = (V, E)$ with vertex set V and edgeset $E \subseteq [V]^2$ a subset $I \subseteq V$ of the vertex set is called *independent*, if the subgraph induced on I contains no edges $e \in E$, i.e., $E \cap [I]^2 = \emptyset$. The maximum cardinality of an independent set I is called the *independence number* $\alpha(\mathcal{G})$ of $\mathcal{G}$. It is well-known, that finding in a graph G an independent set of size $\alpha(G)$ is an NP-hard problem, even for graphs with bounded maximum degree.

This suggests to look for approximation algorithms with guaranteed performance ratio which is the quotient of the sizes of the optimal and the found solution in the worst case. The results of Arora, Lund, Motwani, Sudan and Szegedy [10] on interactive proof systems show that, with respect to polynomial time algorithms, there is no constant performance ratio for the independent set problem for graphs on n vertices, indeed no ratio of $n^{1/4}$ unless $P = NP$, cf. the work of Bellare, Goldreich and Sudan [13]. Recently, Håstad [26] showed that there is no performance ratio of $n^{1/2-\epsilon}$ unless $NP = P$, and no such ratio of $n^{1-\epsilon}$ unless $NP = coR$. With respect to polynomial time algorithms, for triangle-free graphs with maximum degree Δ, a performance ratio of $O(\Delta/\ln\Delta)$ was given in [23], [27], and moreover, if they contain no

*An Extended Abstract of this work appeared in Proc. 8th ACM-SIAM Symposium on Discrete Algorithms SODA'97.

†This research was supported by the Deutsche Forschungsgemeinschaft as part of the Collaborative Research Center "Computational Intelligence" (SFB 531).

complete subgraph K_l, $l \geq 4$, then a performance ratio of $O(\Delta / \ln \ln \Delta)$ is known, cf. [23], [24]. Recently it was shown by Brandt [15] that the independence number of triangle-free graphs with minimum degree $\delta > n/3$ can be computed as fast as matrix multiplication in time $O(n^{2.376})$, while within the class of graphs with minimum degree $\delta > (1-\epsilon)n/4$ where $\epsilon > 0$, the problem is NP-hard. Recall that every triangle-free graph on n vertices has at most $\lfloor n/2 \rfloor \cdot \lceil n/2 \rceil$ edges. For hypergraphs, the corresponding problem has been less studied, cf. [28]. Here we consider this problem of finding a large independent set in a hypergraph. Hypergraphs are important, as many problems can be formulated in terms of them. A *hypergraph* $\mathcal{G} = (V, \mathcal{E})$ is given by a set V of vertices and a set $\mathcal{E}$ of edges (hyperedges) where each edge $E \in \mathcal{E}$ is a nonempty subset of V. A hypergraph $\mathcal{G} = (V, \mathcal{E})$ is called $(k+1)$-*uniform* if $\mathcal{E} \subseteq [V]^{k+1}$, i.e., each edge $E \in \mathcal{E}$ contains exactly $(k+1)$ vertices. If the hypergraph $\mathcal{G} = (V, \mathcal{E})$ is $(k+1)$-uniform, then

$$t(\mathcal{G}) = t = \left(\frac{(k+1) \cdot |\mathcal{E}|}{|V|} \right)^{1/k}$$

is the k'th root of the *average degree* of $\mathcal{G}$. Similarly to the graph case, the *independence number* $\alpha(\mathcal{G})$ of a hypergraph $\mathcal{G} = (V, \mathcal{E})$ is defined as the maximum size of a subset $I \subseteq V$ which contains no edges $E \in \mathcal{E}$, i.e., there is no edge $E \in \mathcal{E}$ with $E \subseteq I$.
We give a general approximation strategy which has a performance ratio of $O(t/(\ln t)^{1/k})$ for $(k+1)$-uniform hypergraphs with average degree t^k and not too many small cycles. For random hypergraphs, no algorithm has a better performance ratio. The idea for doing this originates in a powerful result of Ajtai, Komlós, Pintz, Spencer and Szemerédi [2] on *uncrowded hypergraphs*. These have the property that they contain no small cycles. We remark that derandomizing a probabilistic argument of Spencer [44] yields a performance ratio of $O(t)$ for arbitrary hypergraphs. We apply our approximation strategy to some combinatorial and graph problems for which so far only the existence of solutions of a certain quality was known. Our algorithms match these qualities and yield for these problems algorithmic solutions of state-of-the-art quality.

2 Uncrowded Hypergraphs

Turán's Theorem for Hypergraphs $\mathcal{G}$ gives a lower bound for the independence number $\alpha(\mathcal{G})$, cf. Spencer [44]:

Theorem 2.1 [44] *Let $\mathcal{G} = (V, \mathcal{E})$ be a $(k+1)$-uniform hypergraph, $k \geq 1$, with average degree $t^k \geq 1$ and $|V| = n$. Then,*

$$\alpha(\mathcal{G}) \geq c_k \cdot \frac{n}{t} \,. \tag{1}$$

Moreover, an independent set of size at least $c_k \cdot n/t$ can be found in time $O(|\mathcal{E}|)$.

We remark that for $(k+1)$-uniform hypergraphs on n vertices with average degree $t^k < 1$ one can simply use a Greedy strategy to obtain in time $O(|\mathcal{E}|)$ an independent set of size at least $c'_k \cdot n$, or more simply, we add in an arbitrary way a few edges to the hypergraph to achieve in the resulting hypergraph that $1 \leq t^k < (k+1)$, and then we apply Theorem 2.1.

For completeness we give the algorithmic proof of (1).

Proof: Let $V = \{v_1, v_2, \ldots, v_n\}$ be the set of vertices of $\mathcal{G}$. The existence of an independent set of size as guaranteed in (1) can be shown by picking vertices of V independently of each other at random with probability $p = 1/t$. The set of picked vertices will yield the independent set. For the algorithm we imitate this approach by using the method of conditional probabilities, cf. [9]. To give a deterministic algorithm, we start by generalizing the probabilistic experiment. For $i = 1, 2, \ldots, n$ vertex v_i will be assigned a weight (probability) $p_i \in [0, 1]$. Define a potential $V(p_1, p_2, \ldots, p_n)$ by

$$V(p_1, p_2, \ldots, p_n) = \sum_{i=1}^{n} p_i - \sum_{E \in \mathcal{E}} \prod_{v_i \in E} p_i \ .$$

Note that in the corresponding random experiment, $\sum_{i=1}^{n} p_i$ is the expected number of picked vertices and $\sum_{E \in \mathcal{E}} \prod_{v_i \in E} p_i$ is the expected number of edges in the induced random hypergraph. Now, for $i = 1, 2, \ldots, n$ in each step i we choose either $p_i = 0$ or $p_i = 1$ in order to maximize the value of the potential $V(p_1, p_2, \ldots, p_n)$. As $V(p_1, p_2, \ldots, p_n)$ is linear in each p_i, i.e., for $i = 1$,

$$V(p_1, p_2, \ldots, p_n) = p_1 \cdot V(1, p_2, \ldots, p_n) + (1 - p_1) \cdot V(0, p_2, \ldots, p_n)$$

either $V(1, p_2, \ldots, p_n) \geq V(p_1, p_2, \ldots, p_n)$ or $V(0, p_2, \ldots, p_n) \geq V(p_1, p_2, \ldots, p_n)$. We take vertex v_1 for the independent set, if $V(1, p_2, \ldots, p_n) > V(p_1, p_2, \ldots, p_n)$, else we discard it. Doing this one after the other for $i = 1, 2, \ldots, n$, finally, each vertex v_i is assigned a weight $p_i \in \{0, 1\}$, $i = 1, 2, \ldots, n$.

Choosing in the beginning $p_1 = p_2 = \ldots = p_n = p$, we have

$$V(p, \ldots, p) = p \cdot n - p^{k+1} \cdot \frac{n \cdot t^k}{k+1}$$

which is maximal (taking the derivative) for $p = 1/t$, i.e., with this choice of p we have

$$V(p, \ldots, p) = \frac{k}{k+1} \cdot \frac{n}{t} \ . \tag{2}$$

By our strategy we obtain finally $V(p_1, p_2, \ldots, p_n) \geq V(p, \ldots, p)$. Let $V' = \{v_i \in V \mid p_i = 1\}$. By (2), and using that $p_i \geq 0$ for $i = 1, 2, \ldots, n$, we infer

$$|V'| = \sum_{i=1}^{n} p_i = V(p_1, \ldots, p_n) + \sum_{E \in \mathcal{E}} \prod_{v_i \in E} p_i \geq V(p, \ldots, p) + \sum_{E \in \mathcal{E}} \prod_{v_i \in E} p_i \geq \frac{k}{k+1} \cdot \frac{n}{t} \ .$$

We claim that V' is an independent set in $\mathcal{G}$. Suppose for contradiction that this is not the case and let $E = \{v_{i_1}, v_{i_2}, \ldots, v_{i_{k+1}}\} \in \mathcal{E} \cap [V']^{k+1}$ where $i_1 < i_2 < \ldots < i_{k+1} = s$. Consider step s. As we had chosen in this step $p_s = 1$, we have

$$V(p_1, \ldots, p_{s-1}, 1, p_{s+1}, \ldots, p_n) > V(p_1, \ldots, p_n) \ . \tag{3}$$

However, as edge E came in in step s, we infer

$$V(p_1, \ldots, p_{s-1}, 1, p_{s+1}, \ldots, p_n) - V(p_1, \ldots, p_n) \leq (1 - p_s) - (1 - p_s) = 0$$

which contradicts (3). Thus V' is an independent set in $\mathcal{G}$ of size at least $\frac{k}{k+1} \cdot \frac{n}{t}$, i.e., $\alpha(\mathcal{G}) \geq \frac{k}{k+1} \cdot \frac{n}{t}$.
During the algorithm each vertex and each edge is considered only a constant number of times, hence, with $t^k \geq 1$ the running time is $O(|\mathcal{E}|)$. □

For fixed integers $k \geq 1$, there are examples of $(k+1)$-uniform hypergraphs whose independence numbers match the lower bound (1) up to constant factors. Simply take n/x vertex disjoint copies of complete $(k+1)$-uniform hypergraphs on x vertices each where $t^k \approx \binom{x-1}{k}$. However, quite often the hypergraphs under consideration are in some sense *sparse*, and for these the general lower bound (1) can be improved as can be seen in the following.
A *cycle* of length i in a $(k+1)$-uniform hypergraph $\mathcal{G} = (V, \mathcal{E})$ is a set of i edges from $\mathcal{E}$ whose union contains at most $i \cdot k$ vertices. An *i-cycle* is a cycle of length i which does not contain any cycles of length $2, 3, \ldots, i-1$. To specify the notion 'sparse' mentioned above, we define a hypergraph $\mathcal{G}$ to be *uncrowded* if it does not contain any 2-, 3- or 4-cycles. For a vertex $v \in V$ let $d(v)$ denote its *degree*, i.e., the number of edges $E \in \mathcal{E}$ which contain v. Let $\Delta(\mathcal{G}) = \max_{v \in V} \{d(v)\}$ be the *maximum degree* of $\mathcal{G}$. For a subset $V' \subseteq V$ let $\mathcal{G}' = (V', \mathcal{E}')$ with $\mathcal{E}' = \mathcal{E} \cap [V']^{k+1}$ be the on V' *induced subhypergraph* of $\mathcal{G}$.
For uncrowded hypergraphs, the lower bound (1) was improved by the following powerful result of Ajtai, Komlós, Pintz, Spencer and Szemerédi:

Theorem 2.2 [2] *Let $k \geq 2$ be a fixed integer. Let $\mathcal{G}$ be a $(k+1)$-uniform hypergraph on n vertices. Assume that*

(i) $\mathcal{G}$ is uncrowded, i.e., contains no 2-, 3- or 4-cycles, and

(ii) the maximum degree $\Delta(\mathcal{G})$ satisfies $\Delta(\mathcal{G}) \leq t^k$ where $t \geq t_0(k)$, and

(iii) $n \geq n_0(k, t)$,

then the independence number $\alpha(\mathcal{G})$ satisfies

$$\alpha(\mathcal{G}) \geq \frac{0.98}{e} \cdot 10^{-5/k} \cdot \frac{n}{t} \cdot (\ln t)^{1/k} . \tag{4}$$

Various applications of Theorem 2.2 have been found including the disproof of Heilbronn's conjecture [30], results on Sidon sets [4], Steiner-systems [40], complexity theory [37], Ramsey numbers [3], geometric selection problems [34], Turán numbers for random graphs [29] and graph coloring problems [8], [16] and [33].
Indeed, for a certain range of the involved parameters k, n, t inequality (4) is best possible up to constant factors as a random hypergraph argument shows. Namely, for a fixed integer $k \geq 1$ consider a random $(k+1)$-uniform hypergraph on n vertices, where the edges are chosen independently at random with probability $p = (t/n)^k$, where $n \gg t \gg k$. Let $l = C \cdot n/t \cdot (\ln t)^{1/k}$, where $C > 0$ is a large enough constant. For a fixed l-element subset $I \subseteq V$ of the vertex set, the probability that I is an independent set, is equal to

$$(1-p)^{\binom{l}{k+1}} \leq exp\left\{-p \cdot \binom{l}{k+1}\right\} \leq exp\left\{-\frac{p \cdot l^{k+1} \cdot (1-o(1))}{(k+1)!}\right\}$$

where we used the inequality $1 - x \leq exp\{-x\} := e^{-x}$. Using the inequality $\binom{n}{l} \leq (e \cdot n/l)^l$, the expected number of independent sets of size l in the random hypergraph is at most

$$\begin{aligned}
& \binom{n}{l} \cdot exp\left\{ -\frac{p \cdot l^{k+1} \cdot (1 - o(1))}{(k+1)!} \right\} \\
\leq \quad & \left(\frac{e \cdot n}{l} \cdot exp\left\{ -\frac{p \cdot l^{k} \cdot (1 - o(1))}{(k+1)!} \right\} \right)^l \\
\leq \quad & \left(\frac{e}{C} \cdot \frac{t}{(\ln t)^{1/k}} \cdot exp\left\{ -\frac{C^k \cdot (1 - o(1)) \cdot \ln t}{(k+1)!} \right\} \right)^l \\
\leq \quad & \left(\frac{e}{C} \cdot exp\left\{ \ln t - \frac{\ln \ln t}{k} - \frac{C^k \cdot (1 - o(1)) \cdot \ln t}{(k+1)!} \right\} \right)^l \\
< \quad & 1
\end{aligned}$$

for $C^k \geq 2 \cdot (k+1)!$ and t large enough. However, the expected number of cycles of length $i \leq 4$ in the random hypergraph is at most

$$n^{ik} \cdot p^i = t^{ik} \leq t^{4k} .$$

By Markov's inequality there exists a hypergraph $\mathcal{G}$ on n vertices with at most

$$3 \cdot p \cdot \binom{n}{k+1} \leq 3 \cdot n \cdot \frac{t^k}{(k+1)!}$$

edges which contains no independent set of size $C \cdot n/t \cdot (\ln t)^{1/k}$ and with at most $9 \cdot t^{4k}$ cycles of length at most 4. Note that for each induced subhypergraph $\mathcal{G}'$ of $\mathcal{G}$ we have $\alpha(\mathcal{G}') \leq \alpha(\mathcal{G})$. Now, omitting one vertex from each cycle of length at most 4 gives a $(k+1)$-uniform uncrowded hypergraph on $n - 9 \cdot t^{4k} = (1 - o(1)) \cdot n$ vertices, with average degree at most $3/k! \cdot t^k$ and with independence number at most $C \cdot n/t \cdot (\ln t)^{1/k}$, provided $t^{4k} = o(n)$. Hence, for many hypergraphs inequality (4) is best possible up to constant factors.
Notice, that in (4) we gain a logarithmic factor, if we compare it to Turán's lower bound (1). This additional logarithmic factor is of interest as in several applications one can use Theorem 2.2 to prove lower bounds which, with respect to the corresponding applications, asymptotically match the upper bounds.
Recently, Fundia [19] gave a polynomial time algorithm, which finds an independent set of size asymptotically at least as guaranteed by inequality (4).

Theorem 2.3 [19] *Let $k \geq 2$ be a positive integer. Let $\mathcal{G}$ be an uncrowded $(k+1)$-uniform hypergraph on n vertices with average degree t^k where $k^4 \cdot t^{4k} < n$, and $k < t$. Then, one can find in time $O\left(n^3 \cdot t^{6k} \cdot \ln t\right)$ an independent set of size at least $\Omega\left(n/t \cdot (\ln t)^{1/k}\right)$.*

Notice, that in uncrowded $(k+1)$-uniform hypergraphs $\mathcal{G} = (V, \mathcal{E})$ with $|V| = n$ two edges have at most one vertex in common, hence, $|\mathcal{E}| \leq \binom{n}{2}/\binom{k+1}{2}$.

We remark that in Theorem 2.3 the condition $k^4 \cdot t^{4k} < n$ is not that important, cf. [8], if we ignore for the moment the algorithmic aspects. Namely, if $k^4 \cdot t^{4k} \geq n$, then we set

$$N = \left\lceil \frac{k^4 \cdot t^{4k} + 1}{n} \right\rceil$$

and we form a new hypergraph $\mathcal{H} = (V', \mathcal{E}')$ by taking N vertex disjoint copies of $\mathcal{G}$. Clearly, $|V'| = N \cdot n$ and $|\mathcal{E}'| = N \cdot |\mathcal{E}|$, thus $\mathcal{G}$ and $\mathcal{H}$ have the same average degree. However, $\mathcal{H}$ fulfils the assumptions of Theorem 2.3, and we obtain in time $O\left((N \cdot n)^3 \cdot t^{6k} \cdot \ln t\right)$ an independent set of size $\Omega\left(N \cdot n/t \cdot (\ln t)^{1/k}\right)$. Then, restricting the independent set to one copy of $\mathcal{G}$ yields the desired result, i.e., an independent set of size $\Omega(n/t \cdot (\ln t)^{1/k})$. We have in uncrowded hypergraphs $t = O(n^{1/k})$, hence, we infer for the running time $O\left((N \cdot n)^3 \cdot t^{6k} \cdot \ln t\right) = O\left(t^{18k} \cdot \ln t\right) = O\left(n^{18} \cdot \ln n\right)$, i.e., polynomial running time for fixed k.
Contrary to our considerations of blowing up the hypergraph, i.e., taking copies, we will work on small subhypergraphs of $\mathcal{G}$ to keep the running times of the corresponding algorithms small, cf. Kortsarz and Peleg [31]. Throughout this paper, k will always be a fixed positive integer with $k \geq 2$, and t^k, the average degree, will always be an increasing function of n, the number of vertices of the hypergraph, i.e., $t = t(n) \to \infty$ with $n \to \infty$.
Theorem 2.3 does not seem to be applicable to many hypergraphs. In general, the hypergraphs under consideration are not uncrowded and have quite a lot of 2-, 3- or 4-cycles. However, as we will see, in several cases, the trick is to choose at random an appropriate small subhypergraph which turns out to be 'nearly' uncrowded, i.e., has only a few small cycles. After deleting these cycles one can apply Theorem 2.3. Indeed, seemingly against the intuition, one chooses vertices with a probability $p = t^{-1+\epsilon}$ for some $\epsilon > 0$ which is a little bigger than that one would usually take, i.e., $p = t^{-1}$. But this gives the improvement by a logarithmic factor.
In the following we state our main results. First, we improve Theorem 2.3 as follows.

Theorem 2.4 *Let $k \geq 2$ be a fixed positive integer. Let $\mathcal{G}$ be an uncrowded $(k+1)$-uniform hypergraph on n vertices with average degree t^k and $t \to \infty$ with $n \to \infty$. Then, for each fixed $\delta > 0$, one can find in time*

$$O\left(n \cdot t^k + \frac{n^3}{t^{3-\delta}}\right) \tag{5}$$

an independent set in $\mathcal{G}$ of size at least $\Omega\left(n/t \cdot (\ln t)^{1/k}\right)$.

The parameter $\delta > 0$ can be chosen to be small. Thus, the more edges the hypergraph has, the closer is the running time to the linear one $O(n \cdot t^k)$. However, as $t = O(n^{1/k})$, for $k \geq 3$ and $\delta > 0$, the time bound is always $O(|\mathcal{E}| + n^{3-3/k+\delta}) = O(n^{3-3/k+\delta})$.
The proof of Theorem 2.4 shows that with the factor t^δ in the running time there comes a factor $(\delta/(6k+4))^{1/k}$ in the quality of the solution. Hence, for $k \geq 3$, dropping the running time from $O(n^3/t^{3-\delta})$ to $O(n^3/t^{3-\delta/2})$, i.e., by the factor $t^{\delta/2}$, means that we loose in the guaranteed quality the factor $2^{1/k}$.

Corollary 2.5 *Let $k \geq 2$ be a fixed positive integer. Let $\mathcal{G}$ be an uncrowded $(k+1)$-uniform hypergraph on n vertices with average degree at most t^k and $t \to \infty$ with $n \to \infty$. Then, for each fixed $\delta > 0$, one can find in time*

$$O\left(n \cdot t^k + \frac{n^3}{t^{3-\delta}}\right) \tag{6}$$

an independent set in $\mathcal{G}$ of size at least $\Omega\left(n/t \cdot (\ln t)^{1/k}\right)$.

Proof: If the average degree $t(\mathcal{G})^k$ of $\mathcal{G}$ satisfies $t(\mathcal{G})^k \leq t^k / \ln t$, then by Theorem 2.1 we find in time $O(n \cdot t(\mathcal{G})^k) = O(n \cdot t^k)$ an independent set of size

$$\Omega\left(\frac{n}{t(\mathcal{G})}\right) = \Omega\left(\frac{n}{t} \cdot (\ln t)^{1/k}\right) .$$

Otherwise, if $t^k \geq t(\mathcal{G})^k > t^k / \ln t$, then we apply Theorem 2.4 with $\delta' = \delta/2$ and we obtain in time

$$O\left(n \cdot t(\mathcal{G})^k + \frac{n^3}{t(\mathcal{G})^{3-\delta'}}\right) = O\left(n \cdot t^k + \frac{n^3}{t^{3-\delta'}} \cdot (\ln t)^{(3-\delta')/k}\right) = O\left(n \cdot t^k + \frac{n^3}{t^{3-\delta}}\right)$$

an independent set of size at least $\Omega\left(n/t(\mathcal{G}) \cdot (\ln t(\mathcal{G}))^{1/k}\right) = \Omega\left(n/t \cdot (\ln t)^{1/k}\right)$. Here we used that the function $f(t) = (\ln t)^{1/k}/t$ is decreasing for $t \geq 2$. □

It turns out that the 2-cycles are important. For a hypergraph $\mathcal{G} = (V, \mathcal{E})$ a 2-cycle $\{E_1, E_2\}$ with $E_1, E_2 \in \mathcal{E}$ is called $(2,j)$-*cycle* if $|E_1 \cap E_2| = j$. Let $s_{2,j}(\mathcal{G})$ denote the number of $(2,j)$-cycles in $\mathcal{G}$.

The next theorem yields an algorithmic version of the existence result which was proved by Duke, Lefmann and Rödl [16]. The following corollary 2.8 gives an algorithmic solution of a conjecture of Spencer [45].

Theorem 2.6 *Let $k \geq 2$ be a fixed integer. Let $\mathcal{G} = (V, \mathcal{E})$ be a $(k+1)$-uniform hypergraph on n vertices with average degree t^k where $t \to \infty$ with $n \to \infty$. If the $(2,j)$-cycles satisfy*

$$s_{2,j}(\mathcal{G}) \leq c \cdot n \cdot t^{2k+1-j-\gamma}$$

for $j = 2, 3, \ldots, k$, and some constants $c, \gamma > 0$, then one can find for every fixed $\delta > 0$ in time $O\left(n \cdot t^k + \sum_{j=2}^k s_{2,j}(\mathcal{G}) + n^3/t^{3-\delta}\right)$ an independent set of size at least

$$C(k, \gamma, \delta, c) \cdot \frac{n}{t} \cdot (\ln t)^{1/k} .$$

Corollary 2.7 *Let $k \geq 2$ be a fixed integer. Let $\mathcal{G} = (V, \mathcal{E})$ be a $(k+1)$-uniform hypergraph on n vertices with average degree at most t^k where $t \to \infty$ with $n \to \infty$. If the $(2,j)$-cycles satisfy*

$$s_{2,j}(\mathcal{G}) \leq c \cdot n \cdot t^{2k+1-j-\gamma}$$

for $j = 2, 3, \ldots, k$, and some constants $c, \gamma > 0$, then one can find for every fixed $\delta > 0$ in time $O\left(n \cdot t^k + \sum_{j=2}^k s_{2,j}(\mathcal{G}) + n^3/t^{3-\delta}\right)$ an independent set of size at least $\Omega\left(n/t \cdot (\ln t)^{1/k}\right)$.

Proof: If the average degree $t(\mathcal{G})$ of $\mathcal{G}$ satisfies $t^k \geq t(\mathcal{G})^k \geq t^k/\ln t$, then the $(2,j)$-cycles satisfy

$$s_{2,j}(\mathcal{G}) \leq c \cdot n \cdot t^{2k+1-j-\gamma} \leq c \cdot n \cdot t(\mathcal{G})^{2k+1-j-\gamma/2}$$

for t large. Thus, the assumptions of Theorem 2.6 are fulfilled, and we obtain for every fixed $\delta > 0$ in time $O\left(n \cdot t^k + \sum_{j=2}^k s_{2,j}(\mathcal{G}) + n^3/t^{3-\delta}\right)$ an independent set of size at least $\Omega\left(n/t \cdot (\ln t)^{1/k}\right)$. Otherwise, if $t(\mathcal{G})^k < t^k/\ln t$, we apply Theorem 2.1 and we obtain in time $O(n \cdot t^k)$ an independent set of size at least $\Omega\left(n/t \cdot (\ln t)^{1/k}\right)$. □

As an immediate consequence of Corollary 2.7 we have the following:

Corollary 2.8 *Let $k \geq 2$ be a fixed integer. Let $\mathcal{G} = (V, \mathcal{E})$ be a $(k+1)$-uniform hypergraph on n vertices with average degree at most t^k where $t \to \infty$ with $n \to \infty$. If $\mathcal{G}$ does not contain any 2-cycles, then one can find for every fixed $\delta > 0$ in time $O\left(n \cdot t^k + n^3/t^{3-\delta}\right)$ an independent set of size at least $\Omega\left(n/t \cdot (\ln t)^{1/k}\right)$.*

3 Choosing Small Subhypergraphs

For proving our results, we use the idea of choosing small subhypergraphs on which we control certain parameters like the number of vertices, edges or cycles. The control on the small subhypergraphs reflects a probabilistic approach.

Lemma 3.1 *Let $k \geq 2$ be a positive integer. Let $\mathcal{G} = (V, \mathcal{E})$ be a $(k+1)$-uniform hypergraph with $s_{2,j}(\mathcal{G})$ many $(2,j)$-cycles, $j = 2, 3, \ldots, k$. For every p with $0 \leq p \leq 1$, one can find in time $O(|V| + |\mathcal{E}| + \sum_{j=2}^k s_{2,j}(\mathcal{G}))$ an induced subhypergraph $\mathcal{H} = (V', \mathcal{E}')$ such that for $j = 2, 3, \ldots, k$ it is*

$$|V'| \geq p/3 \cdot |V| \quad \text{and} \quad |\mathcal{E}'| \leq 3 \cdot p^{k+1} \cdot |\mathcal{E}| \quad \text{and} \quad s_{2,j}(\mathcal{H}) \leq 3 \cdot (k-1) \cdot p^{2k+2-j} \cdot s_{2,j}(\mathcal{G}) .$$

Proof: By inspecting each two-element subset which is contained in an edge $E \in \mathcal{E}$, we obtain for each two-element set $\{x, y\}$ all edges $E \in \mathcal{E}$ with $\{x, y\} \subset E$. This can be done in time $O(|\mathcal{E}|)$. Then, the sets C_j, which contain the vertex sets of all $(2,j)$-cycles in $\mathcal{G}$, $j = 2, 3, \ldots, k$, can be constructed in time $O(|V| + |\mathcal{E}| + \sum_{j=2}^k s_{2,j}(\mathcal{G}))$.
Let $V = \{v_1, v_2, \ldots, v_n\}$. Every vertex v_i will be assigned a weight $p_i \in [0,1]$, $i = 1, 2, \ldots, n$. Define a potential $V(p_1, p_2, \ldots, p_n)$ by

$$V(p_1, p_2, \ldots, p_n) = 3^{pn/3} \cdot \prod_{i=1}^n \left(1 - \frac{2}{3} \cdot p_i\right) + \frac{\sum_{E \in \mathcal{E}} \prod_{v_i \in E} p_i}{3 \cdot p^{k+1} \cdot |\mathcal{E}|} + \sum_{j=2}^k \frac{\sum_{C \in C_j} \prod_{v_i \in C} p_i}{3 \cdot (k-1) \cdot p^{2k+2-j} \cdot |C_j|} .$$

If $pn < 6$ any two vertices will do, since they do not yield an edge. Hence, we assume that $pn \geq 6$. With $p_1 = \ldots = p_n = p \geq 6/n$ in the beginning, and using $e^{-x} \geq 1 - x$, we infer

$$\begin{aligned} V(p, \ldots, p) &= 3^{pn/3} \cdot \left(1 - \frac{2}{3} \cdot p\right)^n + \frac{p^{k+1} \cdot |\mathcal{E}|}{3 \cdot p^{k+1} \cdot |\mathcal{E}|} + \sum_{j=2}^k \frac{p^{2k+2-j} \cdot s_j(\mathcal{G})}{3 \cdot (k-1) \cdot p^{2k+2-j} \cdot s_{2,j}(\mathcal{G})} \\ &< \left(\frac{3}{e^2}\right)^{pn/3} + \frac{2}{3} < 1 . \end{aligned}$$

The potential $V(p_1, p_2, \ldots, p_n)$ is linear in each p_i. As in the algorithmic argument for proving Theorem 2.1, step by step, we decide which choice of p_i, either $p_i = 0$ or $p_i = 1$, minimizes the current value of $V(p_1, p_2, \ldots, p_n)$. We choose a vertex v_1 to be in V' iff $V(1, p_2, \ldots, p_n) \leq V(0, p_2, \ldots, p_n)$. Doing this for all vertices $v_1, v_2, \ldots, v_n$ yields the desired set $V' = \{v_i \in V \mid p_i = 1\}$. We always choose the value of p_i to minimize the value of the potential, i.e., finally, we have $V(p_1, p_2, \ldots, p_n) < 1$, too. All summands in $V(p_1, p_2, \ldots, p_n)$ are nonnegative. If the chosen subhypergraph $\mathcal{H} = (V', \mathcal{E}')$ with $\mathcal{E}' = \mathcal{E} \cap [V']^{k+1}$ violated one of the desired properties, then the corresponding summand would be bigger than 1 which is not possible by our strategy.
The computation of $V(p, p, \ldots, p)$ in the beginning can be done in time $O(|V| + |\mathcal{E}| + \sum_{j=2}^{k} s_{2,j}(\mathcal{G}))$, and updating all the potentials $V(p_1, \ldots, p_n)$ during the algorithm can be done in the same time. $\square$

For a hypergraph $\mathcal{G}$ let $\mu_i(\mathcal{G})$, $i = 3, 4$, denote the number of i-cycles in $\mathcal{G}$.

Lemma 3.2 *Let $k \geq 2$ be a positive integer. Let $\mathcal{G} = (V, \mathcal{E})$ be a $(k+1)$-uniform hypergraph on $|V| = n$ vertices with average degree t^k. If $\mathcal{G}$ does not contain any 2-cycles, then one can find in time $O(n + n \cdot t^k)$ an induced subhypergraph $\mathcal{G}' = (V', \mathcal{E}')$ with $|V'| \geq |V|/2$ and*

$$\mu_i(\mathcal{G}) \leq c_i(k) \cdot n \cdot t^{(i-1)k} \tag{7}$$

for $i = 3, 4$. The sets C_i of i-cycles, $i = 3, 4$, can be computed in time $O(n + n \cdot t^{(i-1)k}) = O(n + n \cdot t^{3k})$.

Proof: For each edge $E \in \mathcal{E}$ we mark in time $O(|\mathcal{E}|)$ all two-element subsets which are contained in E by 'E'. Then, in time $O(1)$ we can decide whether a two-element set is contained in some edge $E \in \mathcal{E}$ or not.
First we discard all vertices $v \in V$ with degree $d(v) > 2(k+1) \cdot t^k = \Delta$ and all edges incident with such vertices v. This can be done in time $O(n + n \cdot t^k)$. The resulting induced hypergraph $\mathcal{G}' = (V', \mathcal{E}')$ of $\mathcal{G}$ has $|V'| \geq n/2$ vertices. As $\mathcal{G}$ contains no 2-cycles, for each two distinct vertices $x, y \in V'$, there exists at most one edge which contains both x and y. The maximum degree of $\mathcal{G}'$ is at most Δ, hence

$$\begin{aligned}
\mu_3(\mathcal{G}') &\leq \frac{1}{3} \cdot n \cdot \binom{\Delta}{2} \cdot k^2 \leq c_3(k) \cdot n \cdot t^{2k} \\
\mu_4(\mathcal{G}') &\leq \frac{1}{4} \cdot n \cdot \binom{\Delta}{2} \cdot k \cdot \Delta \cdot k^2 \leq c_4(k) \cdot n \cdot t^{3k} .
\end{aligned}$$

To determine the set C_3 of all 3-cycles in $\mathcal{G}'$, we consider each vertex $v \in V_0$, and all pairs $\{E_1, E_2\} \in [\mathcal{E}']^2$ of edges with $v \in E_1$ and $v \in E_2$. This can be done in time $O(n \cdot \Delta^2)$. Then, for each two vertices $x \in E_1 \setminus \{v\}$ and $y \in E_2 \setminus \{v\}$ we test in time $O(1)$, whether there exists an edge $E \in \mathcal{E}'$ with $x, y \in E$. Thus, determining the set C_3 of 3-cycles in $\mathcal{G}'$ can be done in time $O(n + n \cdot t^{2k})$. Similarly, one can determine the set C_4 of 4-cycles in $\mathcal{G}'$ in time $O(n + n \cdot t^{3k})$. $\square$

Lemma 3.3 *Let $\mathcal{G} = (V, \mathcal{E})$ be a $(k+1)$-uniform hypergraph on $|V| = n$ vertices with average degree t^k where $\mathcal{G}$ contains no 2-cycles. For every p with $0 \le p \le 1$, one can find in time $O(n + n \cdot t^{3k})$ an induced subhypergraph $\mathcal{H} = (V', \mathcal{E}')$ of $\mathcal{G}$ such that*

$$\begin{aligned} |V'| &\geq p/6 \cdot |V| \\ |\mathcal{E}'| &\leq 3 \cdot p^{k+1} \cdot |\mathcal{E}| \\ \mu_i(\mathcal{H}) &\leq 6 \cdot p^{ik} \cdot c_i(k) \cdot n \cdot t^{(i-1)k} \quad \textit{for } i = 3, 4 \ . \end{aligned} \tag{8}$$

Proof: By Lemma 3.2, we find in $\mathcal{G}$ in time $O(n + n \cdot t^k)$ an induced subhypergraph $\mathcal{G}' = (V', \mathcal{E}')$ with $V' = \{v_1, v_2, \ldots, v_l\}$ and $|V'| \ge |V|/2$ which contains at most $c_i(k) \cdot n \cdot t^{(i-1)k}$ i-cycles. The sets C_i' of i-cycles in $\mathcal{G}'$, $i = 3, 4$, can be computed in time $O(n + n \cdot t^{3k})$. Then we apply to $\mathcal{G}'$ the same derandomization technique as in the proof of Lemma 3.1 using the potential

$$V(p_1, p_2, \ldots, p_l) = 3^{pn/6} \prod_{v_i \in V'} \left(1 - \frac{2}{3} \cdot p_i\right) + \frac{\sum_{E \in \mathcal{E}'} \prod_{v_i \in E}}{3 \cdot p^{k+1} \cdot |\mathcal{E}|} + \sum_{j=3}^{4} \frac{\sum_{C \in C_j'} \prod_{v_i \in C} p_i}{6 \cdot p^{jk} \cdot c_j(k) \cdot n \cdot t^{(j-1)k}} \ ,$$

and we obtain in time $O(n + n \cdot t^{3k})$ an induced subhypergraph fulfilling all properties (8). □

Similarly, one can show the following:

Lemma 3.4 *Let $\mathcal{G} = (V, \mathcal{E})$ be a $(k+1)$-uniform hypergraph with $|V| = n$ and with average degree t^k. For every p with $0 \le p \le 1$, one can find in time $O(n + n \cdot t^k)$ an induced subhypergraph $\mathcal{H} = (V', \mathcal{E}')$ such that*

$$|V'| \ge p/2 \cdot |V| \quad \textit{and} \quad |\mathcal{E}'| \le 2 \cdot p^{k+1} \cdot |\mathcal{E}| \ .$$

Proof: We apply the same derandomization technique as in the proof of Lemma 3.1 by using the potential

$$V(p_1, p_2, \ldots, p_n) = 2^{pn/2} \cdot \prod_{i=1}^{n} \left(1 - \frac{p_i}{2}\right) + \frac{\sum_{E \in \mathcal{E}} \prod_{v_i \in E} p_i}{2 \cdot p^{k+1} \cdot |\mathcal{E}|} \ .$$

□

4 Avoiding 2-Cycles

Here we will give the proof of Theorem 2.4.

Proof: Let $\mathcal{G}$ be an uncrowded $(k+1)$-uniform hypergraph on n vertices with average degree t^k, thus $t = O(n^{1/k})$. The idea is, to choose a small induced subhypergraph $\mathcal{G}_0$ of $\mathcal{G}$ to which we apply Theorem 2.3.

First, we apply Lemma 3.4 to $\mathcal{G}$ with $p = t^{-1+\epsilon}$ where $\epsilon = min\left\{\frac{\delta}{6k+4}, \frac{k-1}{4k}\right\}$, and we obtain in time $O(n \cdot t^k)$ an induced subhypergraph $\mathcal{G}_0 = (V_0, \mathcal{E}_0)$ of $\mathcal{G}$ with

$$|V_0| \ge p/2 \cdot |V| \quad \text{and} \quad |\mathcal{E}_0| \le 2 \cdot p^{k+1} \cdot |\mathcal{E}| \ .$$

Notice that if we had chosen the probability to be $p = 1/t$, then the resulting subhypergraph would only have n/t vertices, but our aim is to find an independent set of size at least $\Omega(n/t \cdot (\ln t)^{1/k})$.

The hypergraph $\mathcal{G}_0$ has average degree

$$t(\mathcal{G}_0)^k \leq t_0^k = 4 \cdot (p \cdot t)^k \ .$$

If $t(\mathcal{G}_0) \leq p \cdot t/(\ln(p \cdot t))^{1/k}$, then we apply Theorem 2.1 and we obtain in time $(p \cdot n + p^{k+1} \cdot |\mathcal{E}|)$ an independent set of size

$$O\left(\frac{p \cdot n}{t(\mathcal{G}_0)}\right) = \Omega\left(\frac{p \cdot n}{p \cdot t} \cdot (\ln(p \cdot t))^{1/k}\right) = \Omega\left(\frac{n}{t} \cdot (\ln t)^{1/k}\right) \ .$$

Otherwise, let $t(\mathcal{G}_0) > p \cdot t/(\ln(p \cdot t))^{1/k}$. We have $t^\epsilon/(\epsilon \cdot \ln t)^{1/k} > k$, as k, ϵ are constants and $t \to \infty$ with $n \to \infty$. With $t = O(n^{1/k})$ and $\epsilon \leq (k-1)/4k$ we infer

$$p/2 \cdot n = n/2 \cdot t^{-1+\epsilon} > k^4 \cdot 4^4 \cdot t^{4k\epsilon} = k^4 \cdot t_0^{4k} \ ,$$

hence the assumptions of Theorem 2.3 are fulfilled, and we apply it to the hypergraph $\mathcal{G}_0$. By omitting some more vertices we can assume w.l.o.g. that $|V_0| = p/2 \cdot n$. As $\epsilon \leq \delta/(6k+4)$, we obtain in time

$$O\left((p \cdot n)^3 \cdot (p \cdot t)^{6k} \cdot \ln t\right) = O\left(\frac{n^3}{t^{3-3\epsilon-6k\epsilon}} \cdot \ln t\right) = O\left(\frac{n^3}{t^{3-\delta}}\right)$$

an independent set of size at least

$$\Omega\left(\frac{p \cdot n}{t_0} \cdot (\ln t_0)^{1/k}\right) = \Omega\left(\frac{p \cdot n}{p \cdot t} \cdot (\ln(p \cdot t))^{1/k}\right) = \Omega\left(\frac{n}{t} \cdot (\ln t^\epsilon)^{1/k}\right) = \Omega\left(\frac{n}{t} \cdot (\ln t)^{1/k}\right)$$

as $\epsilon > 0$ is a constant. Here, we used that the function $f(t) = (\ln t)^{1/k}/t$ is decreasing for $t \geq 2$. The algorithm has the desired running time $O\left(n \cdot t^k + n^3/t^{3-\delta}\right)$. □

Indeed, provided an algorithm does not require any assumptions on the mutual growth of k, t, n, we have the following *Meta-Algorithm*:

Let $k \geq 2$ be a fixed integer. Let $\mathcal{G}$ be an uncrowded $(k+1)$-uniform hypergraph on n vertices with average degree t^k and $t \to \infty$ with $n \to \infty$. Let $\delta > 0$ be any fixed real number.
If one can find in time $f(n,t)$ an independent set in $\mathcal{G}$ of size $\Omega(n/t \cdot (\ln t)^{1/k})$, then one can find in time $O(f(n/t^{1-\delta}, t^\delta) + n \cdot t^k)$ an independent set in $\mathcal{G}$ of size $\Omega(n/t \cdot (\ln t)^{1/k})$.

Now we come to the proof of Theorem 2.6.

Proof: Let the $(2,j)$-cycles of $\mathcal{G}$ satisfy $s_{2,j}(\mathcal{G}) \leq c \cdot n \cdot t^{2k+1-j-\gamma}$ for $j = 2, 3, \ldots, k$ and constants $c, \gamma > 0$. Set

$$p = t^{-1+\epsilon} \tag{9}$$

with

$$\epsilon = \min\left\{\frac{\gamma}{2k}, \frac{1}{3k+2}\right\} \ .$$

By Lemma 3.1 we obtain in time

$$O(|V| + |\mathcal{E}| + \sum_{j=2}^{k} s_{2,j}(\mathcal{G})) \tag{10}$$

an induced subhypergraph $\mathcal{G}_0 = (V_0, \mathcal{E}_0)$ of $\mathcal{G}$ such that, possibly after omitting some more vertices, we have

$$\begin{aligned}
|V_0| &= p/3 \cdot |V| \text{ , and} \\
|\mathcal{E}_0| &\leq 3 \cdot p^{k+1} \cdot |\mathcal{E}| \text{ , and} \\
s_{2,j}(\mathcal{G}_0) &\leq 3 \cdot (k-1) \cdot p^{2k+2-j} \cdot s_{2,j}(\mathcal{G}) \qquad \text{for } j = 2, 3, \ldots, k.
\end{aligned}$$

With (9) for $t^{\epsilon} \geq 2$, i.e., $t \geq t_0(\epsilon)$, we have for $j = 2, 3, \ldots, k$ that

$$p^{2k+2-j} \cdot n \cdot t^{2k+1-j-\gamma} \geq 2 \cdot p^{2k+2-j-1} \cdot n \cdot t^{2k+1-j-1-\gamma} \ ,$$

hence, since $\epsilon \leq \gamma/2k$ and $t \geq t_0(k, \gamma)$, i.e., $t^{\gamma} \geq (36 \cdot c \cdot (k-1))^{2k}$, we have

$$\begin{aligned}
\sum_{j=2}^{k} s_{2,j}(\mathcal{G}_0) &\leq \sum_{j=2}^{k} 3 \cdot (k-1) \cdot p^{2k+2-j} \cdot s_{2,j}(\mathcal{G}) \\
&\leq 6 \cdot (k-1) \cdot c \cdot p^{2k} \cdot n \cdot t^{2k-1-\gamma} \\
&\leq 6 \cdot (k-1) \cdot c \cdot n \cdot t^{-1-\gamma+2k\epsilon} \\
&\leq 36 \cdot (k-1) \cdot c \cdot t^{-\frac{\gamma}{2k}} \cdot (p \cdot n/6) \\
&\leq p \cdot n/6 \ ,
\end{aligned}$$

thus,

$$|V_0| \geq 2 \cdot \sum_{j=2}^{k} s_{2,j}(\mathcal{G}_0) \ . \tag{11}$$

The sets of $(2,j)$-cycles in $\mathcal{G}_0$ can be constructed in time $O(p \cdot n + |\mathcal{E}_0| + \sum_{j=2}^{k} s_{2,j}(\mathcal{G}_0))$. By loosing at most half of the vertices, cf. (11), we delete in $\mathcal{G}_0$ in time

$$O\left(p \cdot n + p^{k+1} \cdot |\mathcal{E}| + \sum_{j=2}^{k} s_{2,j}(\mathcal{G}) \cdot p^{2k+2-j}\right) = O\left(\frac{n}{t^{1-\epsilon}} + \frac{n}{t^{1-(k+1)\epsilon}} + \frac{n}{t^{1+\gamma-2k\epsilon}}\right) = o(n)$$

for $\epsilon \leq 1/(k+2)$ and $\epsilon \leq \gamma/2k$ one vertex from each 2-cycle. By deleting possibly some more vertices we obtain an induced subhypergraph $\mathcal{G}_1 = (V_1, \mathcal{E}_1)$ of $\mathcal{G}_0$ with

$$|V_1| = p \cdot n/6 \quad \text{and} \quad |\mathcal{E}_1| \leq 3 \cdot p^{k+1} \cdot |\mathcal{E}| \ ,$$

such that $\mathcal{G}_1$ contains no 2-cycles anymore. Then we have

$$t(\mathcal{G}_1)^k \leq \frac{(k+1) \cdot |\mathcal{E}_1|}{p \cdot n/6} = 18 \cdot (p \cdot t)^k \ . \tag{12}$$

Next we apply Lemma 3.3 to the hypergraph $\mathcal{G}_1$ with

$$p_1 = \left(\frac{1}{p \cdot t}\right)^{1-\epsilon_1} = t^{-\epsilon(1-\epsilon_1)}$$

where $\epsilon_1 = (k-1)/4k$, and we obtain a subhypergraph $\mathcal{G}_2 = (V_2, \mathcal{E}_2)$ of $\mathcal{G}_1$, which, after possibly deleting some more vertices, satisfies

$$\begin{aligned} |V_2| &= p_1 p \cdot n/36 \\ |\mathcal{E}_2| &\leq 3 \cdot p_1^{k+1} \cdot |\mathcal{E}_1| \leq 9 \cdot (p_1 p)^{k+1} \cdot |\mathcal{E}| \ , \end{aligned}$$

and, by (8), (12), for $i = 3, 4$ also fulfills

$$\mu_i(\mathcal{G}_2) \leq 6 \cdot p_1^{ik} \cdot c_i(k) \cdot \frac{p \cdot n}{6} \cdot t(\mathcal{G}_1)^{(i-1)k} \leq c_i(k) \cdot p_1^{ik} \cdot p \cdot n \cdot (18 \cdot p^k \cdot t^k)^{i-1} \ .$$

This can be done in time $O(p \cdot n + p \cdot n \cdot (p \cdot t)^{3k}) = O(n \cdot t^{-1+\epsilon(3k+1)}) = o(n)$. We claim that

$$|V_2| \geq 4 \cdot (\mu_3(\mathcal{G}_2) + \mu_4(\mathcal{G}_2)) \ . \tag{13}$$

Namely, by our choice of p and p_1, i.e., $p = t^{-1+\epsilon}$ and $p_1 = t^{-\epsilon(1-\epsilon_1)}$, we have

$$\begin{aligned} & \frac{p_1 p \cdot n}{36} \geq 8 \cdot c_i(k) \cdot p_1^{ik} \cdot pn \cdot (18 \cdot p^k \cdot t^k)^{i-1} \\ \Longleftrightarrow \quad & 1 \geq 288 \cdot c_i(k) \cdot p_1^{ik-1} \cdot (18 \cdot p^k \cdot t^k)^{i-1} \\ \Longleftrightarrow \quad & 1 \geq 288 \cdot 18^{i-1} \cdot c_i(k) \cdot t^{-\epsilon(k-1)+\epsilon\epsilon_1(ik-1)} \ . \end{aligned}$$

Thus, for $\epsilon_1 < (k-1)/(4k-1)$ and t large enough, (13) holds, and $\mathcal{G}_2$ contains at least four times as many vertices as 3- and 4-cycles. Moreover, the average degree of $\mathcal{G}_2$ satisfies $t(\mathcal{G}_2)^k = O((p_1 p \cdot t)^k)$. We omit in $\mathcal{G}_2$ all vertices of degree bigger than $2 \cdot (k+1) \cdot t(\mathcal{G}_2)^k$, hence, loosing at most $p_1 p \cdot n/72$ vertices, and then we determine in the resulting induced subhypergraph $\mathcal{G}_2'$ of $\mathcal{G}_2$ the sets of 3- and 4-cycles in time

$$O(p_1 p \cdot n + p_1 p \cdot n \cdot t(\mathcal{G}_2)^{3k}) = O(p_1 p \cdot n \cdot (p_1 p \cdot t)^{3k}) = O(n \cdot t^{-1+\epsilon\epsilon_1(3k+1)}) = o(n)$$

for $\epsilon\epsilon_1 < 1/(3k+1)$. By deleting in time $o(n)$ one vertex from each 3- or 4-cycle from $\mathcal{G}_2'$ we obtain a subhypergraph $\mathcal{G}_3 = (V_3, \mathcal{E}_3)$ of $\mathcal{G}_2$, which does not contain any 2-, 3-, or 4-cycles, i.e., $\mathcal{G}_3$ is uncrowded and satisfies w.l.o.g.

$$|V_3| = |V_2|/2 = p_1 p \cdot n/144 \quad \text{and} \quad |\mathcal{E}_3| \leq 9 \cdot (p_1 p)^{k+1} \cdot |\mathcal{E}| \ ,$$

thus,

$$t(\mathcal{G}_3)^k \leq \frac{(k+1) \cdot 9 \cdot (p_1 p)^{k+1} \cdot |\mathcal{E}|}{p_1 p \cdot n/144} = 1296 \cdot (p_1 p)^k \cdot t^k \ .$$

To the hypergraph $\mathcal{G}_3$ we apply Corollary 2.5, and for fixed $\delta > 0$ we obtain in time

$$O\left(n \cdot t(\mathcal{G}_3)^k + \frac{(p_1 p \cdot n)^3}{(p_1 p \cdot t)^{3-\delta}}\right) = O\left(\frac{n^3}{t^{3-\delta}}\right)$$

an independent set in $\mathcal{G}_3$, and hence in $\mathcal{G}$, of size at least

$$\Omega\left(\frac{p_1 p \cdot n}{p_1 p \cdot t} \cdot (\ln (p_1 p \cdot t))^{1/k}\right) = \Omega\left(\frac{n}{t} \cdot (\ln t)^{1/k}\right)$$

as desired. The overall running time is given by (10). $\square$

5 Applications

In this section, we give applications of our results to some combinatorial problems. Our arguments are guided by the probabilistic existence proofs. However, for computational reasons additional ideas are required.

Definition 5.1 *Let X be a set. A triple system $\mathcal{F} \subseteq [X]^3$ is called a* partial Steiner triple system *if for any two vertices $x_1, x_2 \in X$ there is at most one set $F \in \mathcal{F}$ with $\{x_1, x_2\} \subset F$.*

Hence, every partial Steiner triple system $\mathcal{F}$ is nothing else than a 3-uniform hypergraph which contains no 2-cycles, i.e., $|\mathcal{F}| \leq 1/3 \cdot \binom{n}{2}$.
The next result is essentially a reformulation of Corollary 2.8.

Corollary 5.2 *Let X be an n-element set, and let $\mathcal{F} \subset [X]^3$ be a partial Steiner triple system. Then, one can find in time $O(n^2)$ an independent set $I \subseteq X$ with*

$$|I| \geq c \cdot \sqrt{n \cdot \ln n}\,. \tag{14}$$

We remark that the existence of an independence set of size as in (14) was shown first by Phelps and Rödl [36].
Proof: Since $t^2 = O(n)$, by Corollary 2.8 we find such an independent set of size at least $\Omega(\sqrt{n \cdot \ln n})$ in time $O(n^2)$. □

In the following we consider a generalization of partial Steiner triple systems.

Definition 5.3 *Let h, k be positive integers with $h \leq k$. Let X be an n-element set. A family $\mathcal{F} \subseteq [X]^{k+1}$ of $(k+1)$-element subsets is called* (n,k+1,h)-Steiner system, *if for any two distinct sets $F_1, F_2 \in \mathcal{F}$, it holds $|F_1 \cap F_2| < h$.*

In a $(n, k+1, h)$-Steiner system $\mathcal{F}$ on X, each h-element subset of X is contained in at most one set $F \in \mathcal{F}$, hence, $|\mathcal{F}| \leq \binom{n}{h}/\binom{k+1}{h}$. Note that an $(n, 3, 2)$-Steiner system is a partial Steiner triple system. In every $(n, k+1, 1)$-Steiner system $\mathcal{F}$ distinct sets $F, F' \in \mathcal{F}$ are disjoint, hence, $|\mathcal{F}| = O(n)$ and $\alpha(\mathcal{F}) \geq \frac{k}{k+1} \cdot n$, and such an independent set can be found easily in time $O(n)$. For values $h \geq 2$ we have the following result.

Theorem 5.4 *Let X be an n-element set, and let $\mathcal{F} \subset [X]^{k+1}$ be an $(n, k+1, h)$-Steiner system where $h \geq 2$. Then, for every $\delta > 0$ one can find in time $O(n^h + n^{3-3\cdot(h-1)/k+\delta} + n^{2h-3-(h-1)/k+\delta})$ an independent set $I \subseteq X$ with*

$$|I| \geq c \cdot n^{\frac{k-h+1}{k}} \cdot (\ln n)^{\frac{1}{k}}\,. \tag{15}$$

Indeed, the lower bound in (15) can be written in terms of $|\mathcal{F}|$ as

$$|I| \geq c \cdot \left(\frac{n^{k+1}}{|\mathcal{F}|}\right)^{1/k} \cdot \left(\ln\left(\frac{|\mathcal{F}|}{n}\right)\right)^{1/k},$$

however, we will show only (15).

The existence of an independent set of size at least as given in (15) was proved by Rödl and Siňajová [40]. They also showed by using the Local Lemma of Lovász, cf. [5], that there exists an $(n, k+1, h)$-Steiner system $\mathcal{F}$ with independence number bounded from above by $C \cdot n^{(k-h+1)/k} \cdot (\ln n)^{1/k}$ for some constant $C > 0$. Hence, in general for $|\mathcal{F}| = \Theta(n^h)$ one cannot expect, up to constant factors, any better bound than (15).
Concerning the running times in Theorem 5.4 we have $O(n^2)$ for $h = k = 2$, and $O(n^{3-3/k+\delta})$ for $h = 2$ and $k \geq 3$. For $h \geq 3$ we have the time bound $O(n^{2h-3-(h-1)/k+\delta})$.
Proof: Let $\mathcal{F}$ be an $(n, k+1, h)$-Steiner system on X. The corresponding hypergraph $(X, \mathcal{F})$ is $(k+1)$-uniform with $|\mathcal{F}| \leq \binom{n}{h}/\binom{k+1}{h}$ edges and has average degree

$$t(\mathcal{F})^k \leq t^k = \frac{(k+1) \cdot \binom{n}{h}}{\binom{k+1}{h} \cdot n} \leq c_1 \cdot n^{h-1} .$$

First, we consider the $(2, j)$-cycles in $\mathcal{G} = (X, \mathcal{F})$. For each set $F \in \mathcal{F}$, we take a j-element subset $S \subset F$ and try to extend it to a set $F' \in \mathcal{F}$. This can be done in at most $O(n^{h-j})$ ways. Thus,

$$s_{2,j}(\mathcal{G}) \leq c_j \cdot |\mathcal{F}| \cdot n^{h-j} \leq c_{2,j} \cdot n^{2h-j}$$

for $j = 2, 3, \ldots, h-1$. Moreover, constructing the set of $(2, j)$-cycles can be done in time

$$O\left(|\mathcal{F}| \cdot \binom{|V|}{h-j}\right) .$$

Also for every subset $X' \subseteq X$, the induced subsystem $\mathcal{G}' = (X', \mathcal{F}')$ with $\mathcal{F}' = \mathcal{F} \cap [X']^{k+1}$ we have

$$s_{2,j}(\mathcal{G}') \leq c'_{2,j} \cdot |\mathcal{F}'| \cdot \binom{|X'|}{h-j}$$

for $j = 2, 3, \ldots, h-1$ and the set of $(2, j)$-cycles in $\mathcal{G}'$ can be determined in time $O(|\mathcal{F}'| \cdot |X'|^{h-j})$.
Given $\delta > 0$, by Lemma 3.4, for $p = n^{-1/k+\epsilon}$ with $0 < \epsilon \leq \min\{1/k, \delta/(k+h-1)\}$, we select in time $O(n^h)$ an induced subsystem $(X_0, \mathcal{F}_0)$ of $(X, \mathcal{F})$ which satisfies, after possibly omitting some more vertices,

$$|X_0| = p/2 \cdot |X| \quad \text{and} \quad |\mathcal{F}_0| \leq 2 \cdot p^{k+1} \cdot |\mathcal{F}| .$$

Then

$$t(\mathcal{F}_0)^k \leq t_0^k = 4 \cdot (p \cdot t)^k .$$

For sequences a_n, b_n, $n = 1, 2, \ldots$, let $a_n \ll b_n$ if $a_n/b_n \to 0$ with $n \to \infty$. To apply Corollary 2.7 we want to have for some constant $\gamma > 0$ that

$$\begin{aligned} & p^{k+1} \cdot |\mathcal{F}| \cdot \binom{p \cdot n}{h-j} \ll p \cdot n \cdot (t_0)^{2k+1-j-\gamma} \\ \iff \quad & p^{-(k-h+1)+\gamma} \cdot n^{-(j-1)\cdot\frac{k-h+1}{k}+\gamma\cdot\frac{h-1}{k}} \ll 1 \\ \iff \quad & n^{-(j-2)\cdot\frac{k-h+1}{k}-\epsilon\cdot(k-h+1)+\gamma\cdot\frac{h-2+\epsilon k}{k}} \ll 1 \end{aligned}$$

and the last inequality holds, since $j \geq 2$, for $\gamma < (\epsilon \cdot (k-h+1))/((h-2)/k+\epsilon)$. By Corollary 2.7 we obtain an independent set in $\mathcal{F}_0$, hence in $\mathcal{F}$, of size at least

$$\Omega\left(\frac{p \cdot n}{t_0} \cdot (\ln t_0)^{\frac{1}{k}}\right) = \Omega\left(\frac{p \cdot n}{p \cdot t} \cdot (\ln (p \cdot t))^{\frac{1}{k}}\right) = \Omega\left(n^{\frac{k-h+1}{k}} \cdot (\ln n)^{\frac{1}{k}}\right) .$$

The time for doing this is for any $\delta, \delta^* > 0$ with and $\delta^* \cdot \frac{h-1}{k} < \delta$,

$$\begin{aligned} & O\left(p \cdot n + p^{k+1} \cdot n^h + \sum_{j=2}^{h-1} p^{k+1} \cdot n^h \cdot (p \cdot n)^{h-j} + \frac{(p \cdot n)^3}{(t_0)^{3-\delta^*}}\right) \\ = \; & O\left(p^{k+1} \cdot n^h \cdot (p \cdot n)^{h-2} + n^{3-(3-\delta^*)\cdot \frac{h-1}{k}}\right) \\ = \; & O\left(n^{2h-3-\frac{h-1}{k}+\epsilon(k+h-1)} + n^{3-3(h-1)/k+\delta}\right) \\ = \; & O\left(n^{2h-3-\frac{h-1}{k}+\delta} + n^{3-3(h-1)/k+\delta}\right) . \end{aligned}$$

Thus, the total running time is $O(n^h + n^{3-3(h-1)/k+\delta} + n^{2h-3-(h-1)/k+\delta})$ for each given $\delta > 0$. □

The following problem was considered the first time by Erdös and Guy [18]: Determine the maximum cardinality of a subset X of the $n \times n$-grid such that all mutual Euclidean distances between distinct points of X are distinct. By a Greedy-type argument it was shown in [18] that for every $\epsilon > 0$ such a set X with $|X| \geq c_1 \cdot n^{2/3-\epsilon}$ exists. By a probabilistic argument, this lower bound was improved by Thiele [46] to $|X| \geq c_2 \cdot n^{2/3}/(\ln n)^{1/3}$. Subsequently, in [34] the existence of such a set X with $|X| \geq c_3 \cdot n^{2/3}$ was shown. The problem how to achieve this nonconstructive lower bound remained open. Here we will give such an algorithm. Such problems are somewhat related to problems arising from measuring distances using sonar signals, cf. [17], [21] and [22].

Theorem 5.5 *One can determine in time $O(n^6 \cdot \ln n)$ a subset X of the $n \times n$-grid, such that the distances between distinct points of X are mutually distinct and*

$$|X| \geq c \cdot n^{2/3} . \tag{16}$$

We remark, that by results from number theory currently one can show only the upper bound $|X| \leq c \cdot n/(\ln n)^{1/4}$, cf. [18].

Proof: Let $G_n = \{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\}$ denote the vertex set of the $n \times n$-grid. We form a (nonuniform) hypergraph $\mathcal{G} = (V, \mathcal{E}_3 \cup \mathcal{E}_4)$ on $V = G_n$ with $\mathcal{E}_3 \subseteq [V]^3$ and $\mathcal{E}_4 \subseteq [V]^4$ as follows. Let $d(x, y)$ denote the Euclidean distance between x and y, i.e., $d(x, y) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ for $x = (x_1, x_2)$ and $y = (y_1, y_2)$. For pairwise distinct vertices $x, y, z \in V$ we form a *three-element edge* $E = \{x, y, z\} \in \mathcal{E}_3$ if and only if $d(x, y) = d(x, z)$. Moreover, for pairwise distinct vertices $x^1, x^2, x^3, x^4 \in V$ we form a *four-element edge* $E = \{x^1, x^2, x^3, x^4\} \in \mathcal{E}_4$ if and only if $d(x^1, x^2) = d(x^3, x^4)$.

Our strategy will be to find a large independent set $I \subseteq V$ in the hypergraph $\mathcal{G}$. Clearly, an independent set $I \subseteq V$ is (in the grid) a set with mutual distinct distances. We will find this independent set $I \subseteq V$ again by picking vertices at random, and using derandomization. The usual process of choosing a subhypergraph and controlling the number of $(2, 2)$-cycles right from the beginning would result in a running time of $O\left(n^{8+\frac{c}{\ln \ln n}}\right)$. Therefore, some additional ideas

are involved here to get the time bound $O(n^6 \cdot \ln n)$. To achieve this, we will control the $(2,2)$-cycles at a later stage. We remark that $|\mathcal{E}_3 \cup \mathcal{E}_4| = O(n^6 \cdot \ln n)$ as will be seen later. Thus, our algorithm will be linear in the size of $\mathcal{G}$.
In a preprocessing we compute and store the sets S_d of all solutions of the diophantine equation $x^2 + y^2 = d$ for $d = 1, 2, \ldots, 2(n-1)^2$ with integers x and y. For a fixed value of d this can be done in time $O(\sqrt{d})$ by simply inserting $x = -\lfloor\sqrt{d}\rfloor, -\lfloor\sqrt{d}\rfloor + 1, \ldots, \lfloor\sqrt{d}\rfloor$ in $x^2 + y^2 = d$, and solving this for y. This preprocessing can be done in time

$$O\left(\sum_{d=1}^{2(n-1)^2} \sqrt{d}\right) = O\left(n^3\right) .$$

For positive integers d let $r_2(d)$ denote the number of solutions of the equation $x^2 + y^2 = d$ within the integers. By a result of Wigert, cf. [25], we have

$$\max\left\{r_2(d) \mid d = 1, 2, \ldots, n\right\} \le n^{\frac{c'}{\ln\ln n}} \tag{17}$$

for some constant $c' > 0$, hence,

$$|S_d| = r_2(d) \le n^{\frac{c'}{\ln\ln n}}$$

for $d = 1, 2, \ldots, 2(n-1)^2$. We will use the following identity due to Ramanujan [39]

$$\sum_{d=1}^{m} (r_2(d))^2 = \Theta(m \cdot \ln m) . \tag{18}$$

Given the $n \times n$-grid, we first construct the sets $\mathcal{E}_3$ and $\mathcal{E}_4$ of edges. For $d = 1, 2, \ldots, 2(n-1)^2$ and for each vertex $x_0 \in V$ there are at most $|S_d|$ vertices $x^1, x^2, \ldots, x^l \in V$ with $d(x_0, x^j)^2 = d$, $j = 1, 2, \ldots, l$, and all these vertices can easily be determined by a coordinate transformation $x \to x - x_0$ in time $O(r_2(d))$. Then, $\{x_0, x^i, x^j\}$, $1 \le i < j \le l$ yields a three-element edge in $\mathcal{E}_3$. Hence, using (18) the set $\mathcal{E}_3$ of three-element edges can be constructed in time

$$O\left(n^2 \cdot \sum_{d=1}^{2(n-1)^2} \binom{r_2(d)}{2}\right) = O\left(n^2 \cdot \sum_{d=1}^{2(n-1)^2} (r_2(d))^2\right) = O\left(n^4 \cdot \ln n\right) ,$$

and we also infer

$$|\mathcal{E}_3| \le c_3 \cdot n^4 \cdot \ln n . \tag{19}$$

In a similar fashion we construct the set $\mathcal{E}_4$ of four-element edges. Namely, we pick two distinct vertices $x, y \in V$ and consider the sets $A = \{z \in V \mid d(x,z)^2 = d\}$ and $B = \{w \in V \mid d(y,w)^2 = d\}$ for $d = 1, 2, \ldots, 2(n-1)^2$. Then, $\{x, y, z, w\}$ with $z \in A$, $w \in B$ and $z \neq w$ yields a four-element edge $E \in \mathcal{E}_4$. Using (18), this can be done in time

$$O\left(n^2 \cdot n^2 \cdot \sum_{d=1}^{2(n-1)^2} (r_2(d))^2\right) = O\left(n^6 \cdot \ln n\right) ,$$

and, also

$$|\mathcal{E}_4| \leq c_4 \cdot n^6 \cdot \ln n \; . \tag{20}$$

We want to find again a small, but large enough, uncrowded induced subhypergraph in $\mathcal{G}$. The running time will depend essentially on the number of $(2,j)$-cycles which we have to handle and which are determined by pairs of distinct edges from $\mathcal{E}_4$. Consider only the 4-element edges in $\mathcal{G}$, i.e., $\mathcal{G}|\mathcal{E}_4 = (V, \mathcal{E}_4)$. We show how to construct the sets $C_{2,2}$ and $C_{2,3}$ of $(2,2)$- and $(2,3)$-cycles in $\mathcal{G}|\mathcal{E}_4$, $j = 2,3$. Fix an integer d with $1 \leq d \leq 2(n-1)^2$. We fix a vertex $x \in V$, and we consider the set $S(x;d)$ of all vertices $x^* \in V$ with $d(x,x^*)^2 = d$, i.e., $|S(x;d)| \leq r_2(d)$. We pick two more vertices $y, z \in V \setminus \{x, x^*\}$ and determine the sets $S(y;d)$ and $S(z;d)$. Any pair $\{\{x,x^*,y,y^*\},\{x,x^*,z,z^*\}\}$ with $y^* \in S(y;d)$ and $z^* \in S(z;d)$ yields a $(2,2)$-cycle in $\mathcal{G}$ if the vertices x,x^*,y,y^*,z,z^* are pairwise distinct. Moreover, for an edge $E = \{x^1,x^2,x^3,x^4\} \in \mathcal{E}_4$ with $d(x^1,x^2) = d(x^3,x^4)$ and two distinct vertices $y,z \in V \setminus E$ with $d(x^2,x^4) = d(y,z)$ or with $d(x^2,y) = d(x^4,z)$ the pair $\{\{x^1,x^2,x^3,x^4\},\{x^2,x^4,y,z\}\}$ also yields a $(2,2)$-cycle. Hence, using (17) and (20) the time to construct the set $C_{2,2}$ of $(2,2)$-cycles is given by

$$O\left(|\mathcal{E}_4| \cdot n^2 \cdot max\{r_2(d) \mid 1 \leq d \leq 2(n-1)^2\}\right) = O\left(n^{8+\frac{c_1}{\ln\ln n}} \cdot \ln n\right) = O\left(n^{8+\frac{c}{\ln\ln n}}\right) \tag{21}$$

for some constant $c > 0$. Thus, for some constant $c_{2,2} > 0$, we have

$$s_{2,2}(\mathcal{G}|\mathcal{E}_4) \leq c_{2,2} \cdot n^{8+\frac{c}{\ln\ln n}} \; .$$

To construct the set $C_{2,3}$ of $(2,3)$-cycles in $\mathcal{G}|\mathcal{E}_4$, we pick an edge $E \in \mathcal{E}_4$ and a three-element subset $S \subset E$, which determines a constant number of distances. Then S can be extended in at most $O(\max\{r_2(d) \mid 1 \leq d \leq 2(n-1)^2\})$ ways to an edge $E' \in \mathcal{E}_4$ with $S \subset E'$. Thus, we can construct the set $C_{2,3}$ in $\mathcal{G}|\mathcal{E}_4$ in time

$$O(|\mathcal{E}_4| \cdot \max\{r_2(d) \mid 1 \leq d \leq 2(n-1)^2\}) \; , \tag{22}$$

and also for some constant $c_{2,3} > 0$ we have

$$s_{2,3}(\mathcal{G}|\mathcal{E}_4) \leq c'_{2,3} \cdot n^6 \cdot \ln n \cdot n^{\frac{c'}{\ln\ln}} \leq c_{2,3} \cdot n^{6+\frac{c}{\ln\ln n}} \; .$$

However, we do not construct the sets $C_{2,j}$ of $(2,j)$-cycles right from the beginning. We first determine a small, but big enough induced subhypergraph, where we only control the number of vertices and edges, and in this we will construct the sets of $(2,j)$-cycles, $j = 2,3$.
From (21) we infer that every induced subhypergraph $\mathcal{G}' = (V', \mathcal{E}')$ of the hypergraph $\mathcal{G}|\mathcal{E}_4 = (V, \mathcal{E}_4)$, consisting only of the four-element edges, satisfies

$$s_{2,2}(\mathcal{G}') \;=\; O\left(|\mathcal{E}'| \cdot |V'| \cdot \max\{r_2(d) \mid 1 \leq d \leq 2(n-1)^2\}\right) = O\left(|\mathcal{E}'| \cdot |V'| \cdot n^{\frac{c}{\ln\ln n}}\right) . \tag{23}$$

Similarly, with (22) we infer that

$$s_{2,3}(\mathcal{G}') = O\left(|\mathcal{E}'| \cdot n^{\frac{c}{\ln\ln n}}\right) . \tag{24}$$

Also, the set of $(2,2)$-cycles in $\mathcal{G}'$ can be determined in time $O\left(|\mathcal{E}'| \cdot |V'| \cdot n^{\frac{c}{\ln\ln n}}\right)$, and the set of $(2,3)$-cycles can be constructed in time $O\left(|\mathcal{E}'| \cdot n^{\frac{c}{\ln\ln n}}\right)$.

Lemma 5.6 *Let p be a real number, $0 \leq p \leq 1$. Then, one can construct in time $O(|V| + |\mathcal{E}_3| + |\mathcal{E}_4|)$ a subhypergraph $\mathcal{G}_0 = (V_0, \mathcal{E}_3^0 \cup \mathcal{E}_4^0)$ of $\mathcal{G} = (V, \mathcal{E}_3 \cup \mathcal{E}_4)$ with $\mathcal{E}_i^0 \subseteq \mathcal{E}_i$, $i = 3, 4$, such that $|V_0| = p/3 \cdot |V|$ and $|\mathcal{E}_3^0| \leq 3 \cdot p^3 \cdot |\mathcal{E}_3|$ and $|\mathcal{E}_4^0| \leq 3 \cdot p^4 \cdot |\mathcal{E}_4|$.*

Proof: The proof follows earlier arguments, cf. Lemma 3.4. For $V = \{v_1, v_2, \ldots, v_m\}$, we use the potential

$$V(p_1, p_2, \ldots, p_m) = 3^{pm/3} \cdot \prod_{i=1}^{m} \left(1 - \frac{2}{3} \cdot p_i\right) + \frac{\sum_{E \in \mathcal{E}_3} \prod_{v_i \in E} p_i}{3 \cdot p^3 \cdot |\mathcal{E}_3|} + \frac{\sum_{E \in \mathcal{E}_4} \prod_{v_i \in E} p_i}{3 \cdot p^4 \cdot |\mathcal{E}_4|} .$$

We obtain in time $O(|V| + |\mathcal{E}_3| + |\mathcal{E}_4|)$ a subhypergraph $\mathcal{G}_0 = (V_0, \mathcal{E}_3^0 \cup \mathcal{E}_4^0)$ of $\mathcal{G} = (V, \mathcal{E}_3 \cup \mathcal{E}_4)$ with $|V_0| \geq p/3 \cdot |V|$ and $|\mathcal{E}_3^0| \leq 3 \cdot p^3 \cdot |\mathcal{E}_3|$ and $|\mathcal{E}_4^0| \leq 3 \cdot p^4 \cdot |\mathcal{E}_4|$. By omitting possibly some more vertices we get $|V_0| = p/3 \cdot |V|$. ◻

In time $O(|V| + |\mathcal{E}_3| + |\mathcal{E}_4|) = O(n^6 \cdot \ln n)$ we apply Lemma 5.6 to the hypergraph $\mathcal{G}$ with

$$p = n^{-2/3+\epsilon} \tag{25}$$

where $0 < \epsilon < 2/3$, and we obtain a subhypergraph $\mathcal{G}_0 = (V_0, \mathcal{E}_3^0 \cup \mathcal{E}_4^0)$ of $\mathcal{G}$ with

$$|V_0| = p \cdot n^2/3, \quad |\mathcal{E}_3^0| \leq 3 \cdot p^3 \cdot |\mathcal{E}_3|, \quad |\mathcal{E}_4^0| \leq 3 \cdot p^4 \cdot |\mathcal{E}_4| . \tag{26}$$

Virtually we have no control on the number $s_{2,j}(\mathcal{G}_0|\mathcal{E}_4^0)$ of $(2,j)$-cycles, $j = 2, 3$, in $\mathcal{G}_0|\mathcal{E}_4^0$. However, by (23) we know for $\epsilon < 4/15$ that

$$\begin{aligned}
s_{2,2}(\mathcal{G}_0|\mathcal{E}_4^0) &\leq c'_{2,2} \cdot |\mathcal{E}_4^0| \cdot |V_0| \cdot n^{\frac{c}{\ln \ln n}} \\
&\leq c^*_{2,2} \cdot p^4 \cdot n^6 \cdot \ln n \cdot p \cdot n^2 \cdot n^{\frac{c}{\ln \ln n}} \\
&= c^*_{2,2} \cdot p^5 \cdot n^{8+\frac{c}{\ln \ln n}} \cdot \ln n \\
&\leq c^*_{2,2} \cdot n^{14/3+5\epsilon+\frac{c}{\ln \ln n}} \cdot \ln n \\
&= o(n^6 \cdot \ln n) .
\end{aligned} \tag{27}$$

Moreover, for $\epsilon < 2/3$ we infer from (24) that

$$\begin{aligned}
s_{2,3}(\mathcal{G}_0, \mathcal{E}_4^0) &\leq c'_{2,3} \cdot |\mathcal{E}_4^0| \cdot n^{\frac{c}{\ln \ln n}} \\
&\leq c^*_{2,3} \cdot p^4 \cdot n^{6+\frac{c}{\ln \ln n}} \cdot \ln n \\
&\leq c^*_{2,3} \cdot n^{10/3+4\epsilon+\frac{c}{\ln \ln n}} \cdot \ln n \\
&= o(n^6 \cdot \ln n) .
\end{aligned} \tag{28}$$

The value p is chosen in (25) such that the assumptions of Corollary 2.7 are fulfilled for the hypergraph $\mathcal{G}_0|\mathcal{E}_4^0$, i.e., we choose for p that minimal value, for which the following two conditions hold for some constants $c, \gamma > 0$:

$$\begin{aligned}
s_{2,2}(\mathcal{G}_0|\mathcal{E}_4^0) &\leq c \cdot p \cdot n^2 \cdot t(\mathcal{G}_0|\mathcal{E}_4^0)^{5-\gamma} \\
s_{2,3}(\mathcal{G}_0|\mathcal{E}_4^0) &\leq c \cdot p \cdot n^2 \cdot t(\mathcal{G}_0|\mathcal{E}_4^0)^{4-\gamma} .
\end{aligned}$$

We proceed similarly as in the proof of Theorem 2.6. Namely, we use the following lemma whose proof is along the lines of former statements:

Lemma 5.7 *For every p_1 with $0 \le p_1 \le 1$, one can construct in time*

$$O\left(|V_0| + |\mathcal{E}_3^0| + |\mathcal{E}_4^0| + s_{2,3}(\mathcal{G}|\mathcal{E}_4^0) + s_{2,2}(\mathcal{G}_0|\mathcal{E}_4^0)\right)$$

a subhypergraph $\mathcal{G}_1 = (V_1, \mathcal{E}_3^1 \cup \mathcal{E}_4^1)$ of $\mathcal{G}_0 = (V_0, \mathcal{E}_3^0 \cup \mathcal{E}_4^0)$ with $\mathcal{E}_i^1 \subseteq \mathcal{E}_i^0$, $i = 3, 4$, such that

$$|V_1| = p_1/5 \cdot |V_0| \ , \quad |\mathcal{E}_3^1| \le 5 \cdot p_1^3 \cdot |\mathcal{E}_3^0| \ , \quad |\mathcal{E}_4^1| \le 5 \cdot p_1^4 \cdot |\mathcal{E}_4^0| \ , \quad s_{2,3}(\mathcal{G}_1|\mathcal{E}_4^1) \le 5 \cdot p_1^5 \cdot s_{2,3}\left(\mathcal{G}_0|\mathcal{E}_4^0\right) \ ,$$

and $s_{2,2}\left(\mathcal{G}_1|\mathcal{E}_4^1\right) \le 5 \cdot p_1^6 \cdot s_{2,2}\left(\mathcal{G}_0|\mathcal{E}_4^1\right)$.

Before applying Lemma 5.7 we estimate the corresponding running time, namely using (19), (20), (25), (26), (27) and (28) we obtain

$$O\left(|V_0| + |\mathcal{E}_3^0| + |\mathcal{E}_4^0| + s_{2,3}(\mathcal{G}|\mathcal{E}_4^0) + s_{2,2}(\mathcal{G}_0|\mathcal{E}_4^0)\right) = o(n^6 \cdot \ln n) \ .$$

Let

$$p_1 = \left(\frac{(\ln n)^{1/3}}{p \cdot t}\right)^{1-\epsilon_1} = \left(\frac{1}{n^{2/3+\epsilon}}\right)^{1-\epsilon_1} \ ,$$

where

$$\epsilon_1 = \frac{\epsilon}{5\epsilon + 11/3} \ . \tag{29}$$

By Lemma 5.7 we obtain in time $o\left(n^6 \cdot \ln n\right)$ a subhypergraph $\mathcal{G}_1 = \left(V_1, \mathcal{E}_3^1 \cup \mathcal{E}_4^1\right)$ of $\mathcal{G}_0$ with $\mathcal{E}_i^1 \subseteq \mathcal{E}_i^0$, $i = 3, 4$, where

$$|V_1| = \frac{p_1}{5} \cdot |V_0| = \frac{p_1 p}{15} \cdot n^2 = \frac{1}{15} \cdot n^{2/3 + 2\epsilon_1/3 + \epsilon\epsilon_1} \ ,$$

and,

$$|\mathcal{E}_3^1| \le 5 \cdot p_1^3 \cdot |\mathcal{E}_3^0| \le 15 \cdot p_1^3 \cdot p^3 \cdot |\mathcal{E}_3| \le 15 \cdot c_3 \cdot n^{2\epsilon_1 + 3\epsilon\epsilon_1} \cdot \ln n = o(|V_1|) \ .$$

For the number of $(2,3)$-cycles we have using (28) that

$$\begin{aligned}
s_{2,3}\left(\mathcal{G}_1|\mathcal{E}_4^1\right) &\le 5 \cdot p_1^5 \cdot s_{2,3}\left(\mathcal{G}_0|\mathcal{E}_4^0\right) \\
&\le 15 \cdot c_{2,3} \cdot p_1^5 \cdot n^{10/3 + 4\epsilon + \frac{c}{\ln \ln n}} \cdot \ln n \\
&\le 15 \cdot c_{2,3} \cdot n^{-\epsilon + 10\epsilon_1/3 + 5\epsilon\epsilon_1 + \frac{c}{\ln \ln n}} \cdot \ln n \\
&= o(|V_1|) \ .
\end{aligned}$$

Finally, using (27) we obtain for the number of $(2,2)$-cycles

$$\begin{aligned}
s_{2,2}\left(\mathcal{G}_1|\mathcal{E}_4^1\right) &\le 5 \cdot p_1^6 \cdot s_{2,2}(\mathcal{G}_0|\mathcal{E}_4^0) \\
&\le 5 \cdot c_{2,2}^* \cdot p_1^6 \cdot n^{14/3 + 5\epsilon + \frac{c}{\ln \ln n}} \cdot \ln n \\
&\le 5 \cdot c_{2,2}^* \cdot n^{2/3 - \epsilon + 4\epsilon_1 + 6\epsilon\epsilon_1 + \frac{c}{\ln \ln n}} \cdot \ln n \\
&= o(|V_1|) \ .
\end{aligned}$$

Summarizing, we have $|\mathcal{E}_3^1| + s_{2,3}\,(\mathcal{G}_1|\mathcal{E}_4^1) + s_{2,2}\,(\mathcal{G}_1|\mathcal{E}_4^1) = o(|V_1|)$. We delete from $\mathcal{G}_1$ one vertex from each three-element edge $E \in \mathcal{E}_3^1$ and from each $(2,2)$- or $(2,3)$-cycle in $\mathcal{G}_1|\mathcal{E}_4^0$, and, possibly deleting some more vertices, we obtain a 4-uniform induced subhypergraph $\mathcal{G}_2 = (V_2, \mathcal{E}_4^2)$ of $\mathcal{G}_1$ with

$$|V_2| = \frac{|V_1|}{2} = \frac{p_1 p}{30} \cdot n^2 \ ,$$

and

$$|\mathcal{E}_4^2| \leq 15 \cdot c_4 \cdot (p_1 p)^4 \cdot n^6 \cdot \ln n \ ,$$

and $\mathcal{G}_2$ contains no 2-cycles anymore. By Corollary 2.8 we obtain an independent set in $\mathcal{G}_2$, and hence in $\mathcal{G}$, of size

$$\begin{aligned}
\Omega\left(\frac{p_1 p \cdot n^2}{p_1 p \cdot t} \cdot (\ln(p_1 p \cdot t))^{1/3}\right) &= \Omega\left(\frac{n^2}{t} \cdot \left(\ln\left(n^{2\epsilon_1/3+\epsilon\epsilon_1} \cdot \ln n\right)\right)^{1/3}\right) \\
&= \Omega\left(\frac{n^2}{n^{4/3} \cdot (\ln n)^{1/3}} \cdot (\ln n)^{1/3}\right) \\
&= \Omega\left(n^{2/3}\right) \ .
\end{aligned}$$

as $\epsilon, \epsilon_1 > 0$ are fixed.
The time for doing this is for any given δ with $0 < \delta < 1$

$$\begin{aligned}
O\left(p_1 p \cdot n^2 \cdot (p_1 p \cdot t)^3 + \frac{(p_1 p \cdot n^2)^3}{(p_1 p \cdot t)^{3-\delta}}\right) &= O\left((p_1 p)^4 \cdot n^2 \cdot t^3 + \frac{n^6}{t^{3-\delta}}\right) \\
&= O\left((p_1 p)^4 \cdot n^6 \cdot \ln n + n^{2+4\delta/3}\right) \\
&= O\left(n^{2/3+8\epsilon_1/3+4\epsilon\epsilon_1} \cdot \ln n + n^{2+4\delta/3}\right) \\
&= o\left(n^6 \cdot \ln n\right) \ .
\end{aligned}$$

□

The next result gives a k-dimensional version of Theorem 5.5.

Theorem 5.8 *Let $k \geq 3$ be a fixed integer. For every fixed $\delta > 0$, one can find in time $O(n^{6k-22/3+\delta})$ a subset X of the k-dimensional $n \times \ldots \times n$-grid, such that the distances between distinct points of X are mutually distinct and for some constant $c = c(k, \delta) > 0$ it is*

$$|X| \geq c \cdot n^{2/3} \cdot (\ln n)^{1/3} \ . \tag{30}$$

The existence of such a set X which satisfies (30) was shown in [34]. We remark that here only the upper bound $|X| \leq c \cdot \sqrt{k} \cdot n$ is known, as shown by Erdös and Guy [18].
Proof: Let $V = \{1, 2, \ldots, n\}^k$ be the set of vertices of the k-dimensional grid. As in the proof of Theorem 5.5, we form a nonuniform hypergraph $\mathcal{G} = (V, \mathcal{E}_3 \cup \mathcal{E}_4)$ with $\mathcal{E}_3 \subseteq [V]^3$ and $\mathcal{E}_4 \subseteq [V]^4$ as follows. For vertices $x = (x_1, x_2, \ldots, x_k) \in V$ and $y = (y_1, y_2, \ldots, y_k) \in V$ let $d(x, y) = \sqrt{\sum_{i=1}^k (x_i - y_i)^2}$ denote the Euclidean distance between x and y. For pairwise distinct vertices $x, y, z \in V$ let $\{x, y, z\} \in \mathcal{E}_3$ if and only if $d(x, y) = d(x, z)$. Also, for pairwise

distinct vertices $x^1, x^2, x^3, x^4 \in V$ let $\{x^1, x^2, x^3, x^4\} \in \mathcal{E}_4$ if and only if $d(x^1, x^2) = d(x^3, x^4)$. Again our strategy will be to find a large independent set in $\mathcal{G}$. In a preprocessing we compute and store the sets S_d^k of all solutions of the diophantine equation $x_1^2 + x_2^2 + \ldots + x_k^2 = d$ for $d = 1, 2, \ldots, k(n-1)^2$. By inserting integers $x_1, x_2, \ldots, x_{k-1}$ with $|x_i| \leq \sqrt{d}$ and solving for x_k, this can be done in time

$$O\left(\sum_{d=1}^{k(n-1)^2} (2 \cdot \sqrt{d})^{k-1}\right) \quad = \quad O\left(\int_1^{k \cdot n^2} x^{\frac{k-1}{2}} dx\right) = O\left(n^{k+1}\right) \ .$$

For positive integers d, let $r_k(d)$ denote the number of solutions of $x_1^2 + x_2^2 + \ldots + x_k^2 = d$ within the integers. Rewriting this as $x_1^2 + x_2^2 = d - \sum_{i=3}^k x_i^2$, we infer with (17) that

$$r_k(n) \leq (2 \cdot \sqrt{n})^{k-2} \cdot \ \max_{1 \leq d \leq \sqrt{n}} r_2(d) = O\left(n^{\frac{k}{2} - 1 + \frac{c'}{\ln \ln n}}\right) \ .$$

Hence, for fixed integers $k \geq 3$ we have

$$\max\{r_k(d) \mid d = 1, 2, \ldots, k(n-1)^2\} \leq n^{k-2+\frac{c}{\ln \ln n}} \tag{31}$$

for some constant $c > 0$, and, therefore,

$$|S_d^k| = r_k(d) \leq n^{n-2+\frac{c}{\ln \ln n}} \ .$$

For fixed $k \geq 3$, we have by results from [34] that

$$\sum_{d=1}^{n} (r_k(d))^2 = \Theta\left(n^{k-1}\right) \ . \tag{32}$$

First we construct the sets $\mathcal{E}_3$ and $\mathcal{E}_4$. For $d = 1, 2, \ldots, k(n-1)^2$, and for each vertex $x_0 \in V$ there are at most $|S_d^k|$ vertices $x^1, x^2, \ldots, x^l \in V$ with $d(x_0, x^j)^2 = d$, $j = 1, 2, \ldots, l$, and all these vertices can easily be determined by a coordinate transformation in time $O(r_k(d))$. Then, $\{x_0, x^i, x^j\}$, $1 \leq i < j \leq l$ yields a three-element edge $E \in \mathcal{E}_3$. Using (32), the set $\mathcal{E}_3$ of three-element edges can be constructed in time

$$O\left(n^k \cdot \sum_{d=1}^{k(n-1)^2} \binom{r_k(d)}{2}\right) = O\left(n^k \cdot (k \cdot n^2)^{k-1}\right) = O\left(n^{3k-2}\right) \ ,$$

and, also

$$|\mathcal{E}_3| \leq c_3 \cdot n^{3k-2} \ .$$

In a similar fashion we construct the set $\mathcal{E}_4$ of four-element edges. We pick two distinct vertices $x, y \in V$ and consider the sets $A_d = \{z \in V \mid d(x,z)^2 = d\}$ and $B_d = \{w \in V \mid d(y,w)^2 = d\}$ for $d = 1, 2, \ldots, k(n-1)^2$. Then, $\{x, y, z, w\}$ with $z \in A$ and $w \in B$ yields a four-element edge $E \in \mathcal{E}_4$ if x, y, z, w are pairwise distinct. Using (32), this can be done in time

$$O\left(n^k \cdot n^k \cdot \sum_{d=1}^{k(n-1)^2} (r_k(d))^2\right) = O\left(n^{4k-2}\right) \ ,$$

and, also

$$|\mathcal{E}_4| \leq c_4 \cdot n^{4k-2} . \tag{33}$$

As in the proof of Theorem 5.5, and using (31) the set $C_{2,3}$ of $(2,3)$-cycles in the 4-uniform subhypergraph $\mathcal{G}|\mathcal{E}_4 = (V, \mathcal{E}_4)$ can be constructed in time

$$O(|\mathcal{E}_4| \cdot \max\{r_k(d) \mid 1 \leq d \leq k(n-1)^2\}) = O\left(n^{5k-4+\frac{c}{\ln \ln n}}\right) ,$$

and, for some constant $c_{2,3} > 0$, we have

$$s_{2,3}(\mathcal{G}|\mathcal{E}_4) \leq c_{2,3} \cdot n^{5k-4+\frac{c}{\ln \ln n}} .$$

This shows that for each induced subhypergraph $\mathcal{G}' = (V', \mathcal{E}'_4)$ of $\mathcal{G}|\mathcal{E}_4 = (V, \mathcal{E}_4)$ the set of $(2,3)$-cycles can be constructed in time

$$O\left(|\mathcal{E}'_4| \cdot n^{k-2+\frac{c}{\ln \ln n}}\right)$$

and,

$$s_{2,3}(\mathcal{G}') = O\left(|\mathcal{E}'_4| \cdot n^{k-2+\frac{c}{\ln \ln n}}\right) . \tag{34}$$

Again, as in the proof of Theorem 5.5, using (31) and (33), the construction of the set $C_{2,2}$ of $(2,2)$-cycles in $\mathcal{G}|\mathcal{E}_4$ can be done in time

$$O\left(|\mathcal{E}_4| \cdot |V| \cdot n^{k-2+\frac{c}{\log \log n}}\right) = O\left(n^{6k-4+\frac{c}{\ln \ln n}}\right) ,$$

and, for some constant $c_{2,2} > 0$, we have

$$s_{2,2}(\mathcal{G}|\mathcal{E}_4) \leq c_{2,2} \cdot n^{6k-4+\frac{c}{\ln \ln n}} .$$

Moreover, for every subhypergraph $\mathcal{G}' = (V', \mathcal{E}_4^{'})$ of $\mathcal{G}|\mathcal{E}_4 = (V, \mathcal{E}_4)$ we have

$$s_{2,2}(\mathcal{G}') = O\left(|\mathcal{E}'_4| \cdot |V'| \cdot n^{k-2+\frac{c}{\ln \ln n}}\right) . \tag{35}$$

By Lemma 5.6 with $p = n^{-2/3+\epsilon}$ where $0 < \epsilon < 2/3$, we obtain in time $O(|V| + |\mathcal{E}_3| + |\mathcal{E}_4|) = O\left(n^{4k-2}\right)$ a subhypergraph $\mathcal{G}_0 = (V_0, \mathcal{E}_3^0 \cup \mathcal{E}_4^0)$ of $\mathcal{G}$ with $\mathcal{E}_3^0 \subseteq \mathcal{E}_3$ and $\mathcal{E}_4^0 \subseteq \mathcal{E}_4$ such that

$$|V_0| = p/3 \cdot |V|, \qquad |\mathcal{E}_3^0| \leq 3 \cdot p^3 \cdot |\mathcal{E}_3|, \qquad |\mathcal{E}_4^0| \leq 3 \cdot p^4 \cdot |\mathcal{E}_4|.$$

The value of p is chosen such that the assumptions of Corollary 2.7 are fulfilled for some constants $c, \gamma > 0$:

$$\begin{aligned} s_{2,2}(\mathcal{G}_0|\mathcal{E}_4^0) &\leq c \cdot p \cdot n^k \cdot (p \cdot t)^{5-\gamma} \\ s_{2,3}(\mathcal{G}_0|\mathcal{E}_4^0) &\leq c \cdot p \cdot n^k \cdot (p \cdot t)^{4-\gamma} . \end{aligned}$$

Let t^3 be the average degree of $\mathcal{G}_4|\mathcal{E}_4$. By (33) we have

$$t^3 \leq \frac{4 \cdot c_4 \cdot n^{4k-2}}{n^k} \leq 4 \cdot c_4 \cdot n^{3k-2} \ . \tag{36}$$

By Lemma 5.7 with $p_1 = (p \cdot t)^{-1+\epsilon_1}$ where $\epsilon_1 = \epsilon/(5k + 5\epsilon)$, we obtain a subhypergraph $\mathcal{G}_1 = (V_1, \mathcal{E}_3^1 \cup \mathcal{E}_4^1)$ of $\mathcal{G}_0$ with $\mathcal{E}_i^1 \subseteq \mathcal{E}_i^0$ for $i = 3, 4$ such that $|V_1| = p_1/5 \cdot |V_0|$, $|\mathcal{E}_3^1| \leq 5 \cdot p_1^3 \cdot |\mathcal{E}_3^0|$, $|\mathcal{E}_4^1| \leq 5 \cdot p_1^4 \cdot |\mathcal{E}_4^0|$, $s_{2,3}(\mathcal{G}_1|\mathcal{E}_4^1) \leq 5 \cdot s_{2,3}(\mathcal{G}_1|\mathcal{E}_4^1)$, and $s_{2,2}(\mathcal{G}_1|\mathcal{E}_4^1) \leq 5 \cdot s_{2,2}(\mathcal{G}_1|\mathcal{E}_4^1)$. This can be done in time

$$\begin{aligned}
& O(|V_0| + |\mathcal{E}_3^0| + |\mathcal{E}_4^0| + s_{2,3}(\mathcal{G}_0|\mathcal{E}_4^0) + s_{2,2}(\mathcal{G}_0|\mathcal{E}_4^0)) \\
= \ & O\left(p \cdot n^k + p^3 \cdot |\mathcal{E}_3| + p^4 \cdot |\mathcal{E}_4| + p^4 \cdot |\mathcal{E}_4| \cdot n^{k-2+\frac{c}{\ln\ln n}} + p^4 \cdot |\mathcal{E}_4| \cdot p \cdot n^k \cdot n^{k-2+\frac{c}{\ln\ln n}}\right) \\
= \ & O\left(n^{k-2/3+\epsilon} + n^{3k-4+3\epsilon} + n^{6k-22/3+5\epsilon+\frac{c}{\ln\ln n}}\right) \\
= \ & O\left(n^{6k-22/3+5\epsilon+\frac{c}{\ln\ln n}}\right) \ .
\end{aligned} \tag{37}$$

Since $p_1 p = c \cdot n^{-k+2/3+\epsilon_1(k-4/3+\epsilon)}$ we have by choice of ϵ_1 that

$$|\mathcal{E}_3^1| \ \leq \ 15 \cdot (p_1 p)^3 \cdot |\mathcal{E}_3| \leq C_3 \cdot n^{3\epsilon_1(k-4/3+\epsilon)}) = o(|V_1|) \ .$$

Also, by choice of ϵ_1 we have

$$\begin{aligned}
s_{2,2}(\mathcal{G}_1|\mathcal{E}_4^1) &= O(p_1^6 \cdot s_{2,2}(\mathcal{G}_0|\mathcal{E}_4^0) = O(n^{2/3-\epsilon+6\epsilon_1(k-4/3+\epsilon)+\frac{c}{\ln\ln n}}) = o(|V_1|) \\
s_{2,3}(\mathcal{G}_1|\mathcal{E}_4^1) &= O(p_1^5 \cdot s_{2,3}(\mathcal{G}_0|\mathcal{E}_4^0) = O(n^{-2/3+5\epsilon_1(k-4/3+\epsilon)+\frac{c}{\ln\ln n}}) = o(|V_1|) \ .
\end{aligned}$$

In the resulting subhypergraph $\mathcal{G}_1 = (V_1, \mathcal{E}_3^1 \cup \mathcal{E}_4^1)$ we delete in time $O(|V_1|)$ one vertex from each three-element edge and from all (2,2)- and (2,3)-cycles in $\mathcal{G}_1|\mathcal{E}_4^1$. In loosing $o(|V_1|)$ vertices, we obtain a subhypergraph $\mathcal{G}_2 = (V_2, \mathcal{E}_4^2)$ which contains no 2-cycles anymore, and, possibly after deleting some more vertices, satisfies

$$|V_2| = |V_1|/2 = 1/30 \cdot p_1 p \cdot n^k$$

and

$$|\mathcal{E}_4^2| \leq 15 \cdot c_4 \cdot (p_1 p)^4 \cdot n^{4k-2}.$$

We apply Corollary 2.8 to $\mathcal{G}_2$ and we obtain an independent set in $\mathcal{G}_2$, hence in $\mathcal{G}$ of size at least

$$\Omega\left(\frac{p_1 p \cdot n^k}{p_1 p \cdot t} \cdot (\ln(p_1 p \cdot t))^{1/3}\right) = \Omega\left(\frac{n^k}{n^{k-2/3}} \cdot (\ln n)^{1/3}\right) = \Omega\left(n^{2/3} \cdot (\ln n)^{1/3}\right) \ .$$

The time for doing this is for $0 < \delta_1 < 1$ and $\delta_2 = (k - 2/3) \cdot \delta_1$ given by

$$O\left(p_1 p \cdot n^k \cdot (p_1 p \cdot t)^3 + \frac{(p_1 p \cdot n^k)^3}{(p_1 p \cdot t)^{3-\delta_1}}\right) = O\left((p_1 p)^4 \cdot n^k \cdot t^3 + n^{2+\delta_2}\right) = o\left(n^{6k-22/3+5\epsilon+\frac{c}{\ln\ln n}}\right) \ .$$

Hence, for $5\epsilon < \delta$ by (37) the overall running time is $O\left(n^{6k-22/3+5\epsilon+\frac{c}{\ln\ln n}}\right) = O\left(n^{6k-22/3+\delta}\right)$. □

The next problem which we consider concerns colorings of the edges of a complete graph. The colorings are such that each color class is a *matching*, that is, edges of the same color have no vertex in common.

Theorem 5.9 *Let the edges of the complete graph K_n be colored such that every color class is a matching. Then, one can determine in time $O(n^3)$ a complete subgraph K_l of K_n such that the edges of K_l are totally multicolored (injectively colored), and*

$$l \geq c \cdot (n \cdot \ln n)^{1/3} . \tag{38}$$

The existence of such a totally multicolored subgraph satisfying (38) was shown in [8]. Moreover, Babai [11] showed the existence of colorings of the complete graph K_n where every color class is a matching, such that every totally multicolored complete subgraph K_l satisfies $l \leq C \cdot (n \cdot \ln n)^{1/3}$ where $C \approx 7.3$. The value of the constant C was improved in [33] to $C = 2.21$. Hence, up to a constant factor we cannot do better in Theorem 5.9 with respect to the size of l.

Proof: Let $\Delta: E(K_n) \longrightarrow N$ be a coloring of the edges of the complete graph K_n with n-element vertex set V, where every color class is a matching.

We form a 4-uniform hypergraph $\mathcal{G} = (V, \mathcal{E})$ on the same vertex set as K_n by collecting pairs of edges of the same color. For distinct edges $e_1 = \{v_1, w_1\} \in E(K_n)$ and $e_2 = \{v_2, w_2\} \in E(K_n)$ we have $E = e_1 \cup e_2 \in \mathcal{E} \in [V]^4$ if and only if $\Delta(e_1) = \Delta(e_2)$, i.e., $\Delta(e_1) = \Delta(e_2)$ implies $e_1 \cap e_2 = \emptyset$, as every color class is a matching. With $|\Delta^{-1}(i)| \leq n/2$ for $i \in N$ we infer

$$|\mathcal{E}| = \sum_{i \in N} \binom{|\Delta^{-1}(i)|}{2} \leq \frac{\binom{n}{2}}{\frac{n}{2}} \cdot \binom{\frac{n}{2}}{2} < \frac{1}{8} \cdot n^3 . \tag{39}$$

By first sorting the edges in $E(K_n)$ according to their color, the hypergraph $\mathcal{G}$ can be constructed in time $O(n^3)$. The average degree $t(\mathcal{G})^3$ of $\mathcal{G}$ satisfies

$$t(\mathcal{G})^3 = \frac{4 \cdot |\mathcal{E}|}{|V|} \leq \frac{n^2}{2} = t^3 .$$

First, we construct the $(2,3)$-cycles in $\mathcal{G}$. For each edge $E \in \mathcal{E}$ and each three-element subset $S \subset E$, the set S can be extended in at most a constant number of ways to an edge $E' \in \mathcal{E}$, i.e., $S \subset E'$, as every color class is a matching. Thus the set of $(2,3)$-cycles can be constructed in time $O(|\mathcal{E}|) = O(n^3)$, and for some constant $c_{2,3} > 0$ we have

$$s_{2,3}(\mathcal{G}) \leq c_{2,3} \cdot n^3 . \tag{40}$$

We do not construct the set of $(2,2)$-cycles right now. To do this, observe that for each edge $E \in \mathcal{E}$ and each two-element subset $e \in [E]^2$ there are less than n edges $e' \in E(K_n)$ with $\Delta(e) = \Delta(e')$. Also for distinct vertices $v, w \in E$ and $z \in V \setminus E$, there is at most one edge $e' \in E(K_n)$ with $w \in e'$ and $\Delta(\{v,z\}) = \Delta(e')$. Thus the set of $(2,2)$-cycles can be constructed in time $O(|\mathcal{E}| \cdot n) = O(n^4)$. Similarly, for each subhypergraph $\mathcal{G}' = (V', \mathcal{E}')$ of $\mathcal{G}$, its set of $(2,2)$-cycles can be constructed in time $O(|\mathcal{E}'| \cdot |V'|)$, and, also

$$s_{2,2}(\mathcal{G}') \leq c'_{2,2} \cdot |\mathcal{E}'| \cdot |V'| . \tag{41}$$

By Lemma 5.6, for $p = n^{-1/3+\epsilon}$ with $0 < \epsilon < 1/3$, using (39) and (40) we find in time $O(|V| + |\mathcal{E}| + s_{2,3}(\mathcal{G})) = O(n^3)$ a subhypergraph $\mathcal{G}_0 = (V_0, \mathcal{E}_0)$ of $\mathcal{G} = (V, \mathcal{E})$ with

$$|V_0| = p/3 \cdot |V| , \quad |\mathcal{E}_0| \leq 3 \cdot p^4 \cdot |\mathcal{E}| , \quad \text{and} \quad s_{2,3}(\mathcal{G}_0) \leq 3 \cdot p^5 \cdot s_{2,3}(\mathcal{G}) . \tag{42}$$

Then, $t(\mathcal{G}_0)^3 \leq 9 \cdot (p \cdot t)^3$. We claim that for $\gamma = \epsilon/(1+\epsilon)$ the following holds:

$$s_{2,2}(\mathcal{G}_0) \ll p \cdot n \cdot (p \cdot t)^{5-\gamma} .$$

To see this, observe, using (41), that

$$\begin{aligned} & p^4 \cdot |\mathcal{E}| \cdot p \cdot n \ll p \cdot n \cdot (p \cdot t)^{5-\gamma} \\ \iff\ & p^{-1+\gamma} \cdot n^{-1/3+2\gamma/3} \ll 1 \\ \iff\ & n^{-\epsilon+\gamma(\epsilon+1/3)} \ll 1 \end{aligned}$$

and the last inequality holds. Moreover, by (40) and (42) we have

$$s_{2,3}(\mathcal{G}_0) \leq c_{2,3} \cdot 3 \cdot p^5 \cdot s_{2,3}(\mathcal{G}) \leq c'_{2,3} \cdot p^5 \cdot n^3 \ll p \cdot n \cdot (p \cdot t)^{4-\gamma} .$$

The assumptions of Corollary 2.7 are fulfilled and we obtain an independent set of size at least

$$\Omega\left(\frac{p \cdot n}{p \cdot t} \cdot (\ln(p \cdot t))^{1/3}\right) = \Omega\left((n \cdot \ln n)^{1/3}\right) .$$

The time for doing this is by (41) and (42) for $0 < \delta < 1$ and $\epsilon < 2/15$ given by

$$\begin{aligned} & O\left(p \cdot n + p^4 \cdot n^3 + s_{2,2}(\mathcal{G}_0) + s_{2,3}(\mathcal{G}_0) + \frac{(p \cdot n)^3}{(p \cdot t)^{3-\delta}}\right) \\ = \ & O\left(p \cdot n + p^4 \cdot n^3 + p^5 \cdot |\mathcal{E}| \cdot n + p^5 \cdot n^3 + \frac{n^3}{t^{3-\delta}}\right) \\ = \ & O\left(n^{7/3+5\epsilon} + \frac{n^3}{n^{2-2\delta/3}}\right) \\ = \ & O\left(n^{7/3+5\epsilon} + n^{1+2\delta/3}\right) \\ = \ & o(n^3) . \end{aligned}$$

□

Closely related to the problem just considered, is that of finding large Sidon-sets in Abelian groups [11].

Definition 5.10 *Let $(G, +)$ be an Abelian group. A subset $S \subset G$ is called a* Sidon set *if all pairwise sums $s_1 + s_2$ with $s_1, s_2 \in S$ and $s_1 \neq s_2$ are distinct.*

In the following we will assume that each addition $g + h$ of elements $g, h \in G$ can be done in constant time.

Corollary 5.11 *Let $(G, +)$ be an Abelian group. Let $W \subseteq G$ be an n-element subset of G. Then one can compute in time $O(n^3)$ a Sidon-set $S \subseteq W$ with*

$$|S| \geq c \cdot (n \cdot \ln n)^{1/3} . \tag{43}$$

Note that if $W = G$, where $|W| = n$, each Sidon set S satisfies $\binom{|S|}{2} \leq n$, i.e., $|S| \leq 1 + \sqrt{2 \cdot n}$. The existence of a Sidon set with size as in (43) follows from results in [8].
We remark that within the set of integers, finding a Sidon set $S \subseteq \{1, 2, \ldots, n\}$ with $|S| \geq c \cdot \sqrt{n}$ can easily be done in time $O(n)$ using Singer sets, cf. [43].
Proof: Given the set $W \subseteq G$ with $|W| = n$, we form a complete graph K_n with vertex set W. Then we color the edges $\{v, w\}$ of K_n by color $v + w$. This can be done in time $O(n^2)$. As $(G, +)$ is Abelian, every color class is a matching. Then, we apply Theorem 5.9 to our colored K_n, and we obtain in time $O(n^3)$ a totally multicolored complete subgraph K_l with $l \geq c \cdot (n \cdot \ln n)^{1/3}$. As K_l is totally multicolored, the vertices of K_l yield in W a Sidon set. □

The next problem - Sidon sets within the set of squares of integers - was considered the first by Alon and Erdös [7], where for any $\epsilon > 0$ the lower bound $|S| \geq c \cdot n^{2/3 - \epsilon}$ was proved by a Turán strategy.

Theorem 5.12 *For every $\delta > 0$, one can find in time $O(n^{2+\delta})$ a Sidon set $S \subset \{1^2, 2^2, \ldots, n^2\}$ with*

$$|S| \geq c_\delta \cdot n^{2/3} . \tag{44}$$

The existence of an independent set of size at least as in (44) was proved in [34]. By a result of Landau [32], one has the upper bound $|S| \leq c \cdot n/(\ln n)^{1/4}$ for every Sidon set $S \subseteq \{1^2, 2^2, \ldots, n^2\}$.
Proof: We consider a 4-uniform hypergraph $\mathcal{G} = (V, \mathcal{E})$ with vertex set $V = \{1^2, 2^2, \ldots, n^2\}$ and edgeset $\mathcal{E} \subseteq [V]^4$. Let $\{i_1^2, i_2^2, i_3^2, i_4^2\} \in \mathcal{E}$ if and only if $i_1^2 + i_2^2 = i_3^2 + i_4^2$. To construct $\mathcal{G}$, we consider all pairs (i^2, j^2), $1 \leq i < j \leq n$, and sort them according to the value of their sums $i^2 + j^2$ in time $O(n^2 \cdot \ln n)$. Then, we collect pairwise the pairs with the same value of the sum. Recall, that $r_2(d)$ denotes the number of representations of d as a sum of two squares, i.e., $x^2 + y^2 = d$ for integers x, y. By (17), we have

$$r_2(n) \leq n^{\frac{c'}{\ln \ln n}} .$$

Using (18), we infer

$$|\mathcal{E}| \leq \sum_{d=1}^{2n^2} \binom{r_2(d)}{2} \leq c_4 \cdot n^2 \cdot \ln n \tag{45}$$

and $\mathcal{E}$ can be constructed in time $O(n^2 \cdot \ln n)$.
Now we consider the 2-cycles in the hypergraph $\mathcal{G}$. To count the number $s_{2,2}(\mathcal{G}')$ of $(2,2)$-cycles in any subhypergraph $\mathcal{G}' = (V', \mathcal{E}')$ consider a fixed edge $\{i_1^2, i_2^2, i_3^2, i_4^2\} \in \mathcal{E}'$. The number of edges $\{i_1^2, i_2^2, x^2, y^2\} \in \mathcal{E}'$ with $i_1^2 + i_2^2 = x^2 + y^2$ is at most $r_2(i_1^2 + i_2^2) \leq n^{\frac{c_1}{\ln \ln n}}$. Moreover, the number of edges $\{i_1^2, i_2^2, x^2, y^2\} \in \mathcal{E}'$ with $i_1^2 + x^2 = i_2^2 + y^2$ is given by a constant times the number of divisors of $i_1^2 - i_2^2$, hence, is at most $n^{\frac{c_2}{\ln \ln n}}$ cf. [25], [34]. Thus, we have for constants $c_3, c_{2,2} > 0$ that

$$s_{2,2}(\mathcal{G}') \leq c_{2,2} \cdot |\mathcal{E}'| \cdot n^{\frac{c_3}{\ln \ln n}} ,$$

and the set of $(2,2)$-cycles in $\mathcal{G}'$ can be constructed in time $O\left(|\mathcal{E}'| \cdot n^{\frac{c_3}{\ln \ln n}}\right)$.

As every three-element subset $S \subset V$ can be extended only in a constant number of ways to an edge $E \in \mathcal{E}$, i.e., $S \subset E$, every subhypergraph $\mathcal{G}' = (V', \mathcal{E}')$ of $\mathcal{G}$ satisfies for some constant $c_{2,3} > 0$ that

$$s_{2,3}(\mathcal{G}') \leq c_{2,3} \cdot |\mathcal{E}'| ,$$

and the set of $(2,3)$-cycles in $\mathcal{G}'$ can be constructed in time $O(|\mathcal{E}'|)$.
Now, for the average degree $t(\mathcal{G})^3$ of $\mathcal{G}$ we have by (45) that

$$t(\mathcal{G})^3 \leq t^3 = \frac{4 \cdot c_4 \cdot n^2 \cdot \ln n}{n} = c_5 \cdot n \cdot \ln n .$$

Then, for $0 < \gamma < 1/3$, some constant $c > 0$ and n large we have

$$\begin{aligned} s_{2,2}(\mathcal{G}) &\leq c_{2,2} \cdot n^2 \cdot \ln n \cdot n^{\frac{c_3}{\ln \ln n}} \ll n^{8/3 - \gamma/3} \leq c \cdot n \cdot t^{5-\gamma} \\ s_{2,3}(\mathcal{G}) &\leq c_{2,3} \cdot n^2 \cdot \ln n \ll n^{7/3 - \gamma/3} \leq c \cdot n \cdot t^{4-\gamma}, \end{aligned}$$

hence, the assumptions of Corollary 2.7 are fulfilled, and we obtain for any $\delta' > 0$ with $\delta' < 3\delta$ in time

$$O\left(n^2 \cdot \ln n \cdot n^{\frac{c_3}{\ln \ln n}} + \frac{n^3}{t^{3-\delta'}}\right) = O\left(n^{2+\delta'/3} \cdot (\ln n)^{\delta'/3}\right) = O\left(n^{2+\delta}\right)$$

an independent set of size at least

$$\Omega\left(\frac{n}{(n \cdot (\ln n))^{1/3}} \cdot (\ln n)^{1/3}\right) = \Omega\left(n^{2/3}\right) .$$

□

Sidon sets in arbitrary groups were first considered by Babai and Sós [12]. We distinguish two types:

Definition 5.13 *Let $(G, \cdot)$ be a group and let S be a subset of G. The set S is a Sidon set of the* first kind *if for all $x, y, z, w \in S$ where at least three are distinct, it is*

$$x \cdot y \neq z \cdot w .$$

The set S is a Sidon set of the second kind *if*

$$x \cdot y^{-1} \neq z \cdot w^{-1} .$$

Theorem 5.14 *Let $(G, \cdot)$ be an arbitrary group. Then for every subset $W \subseteq G$ with $|W| = n$, one can determine in time $O(n^3)$ a subset $S \subseteq W$, which is a Sidon set of both kinds, first and second, and satisfies*

$$|S| \geq c \cdot (n \cdot \ln n)^{1/3} . \tag{46}$$

The existence of a Sidon set with $|S| \geq c \cdot n^{1/3}$ was shown first by Babai and Sós [12]. In [33] the existence of a Sidon set S satisfying (46) was shown.

Proof: We sketch the arguments for Sidon sets of the second kind. The arguments for Sidon sets of the first kind are similar (also for Sidon sets of both kinds). By results from [12], a subset $S \subseteq G$ is a Sidon set of the second kind if for pairwise distinct elements $x, y, z \in S$ it is

$$(i) \qquad\qquad x \cdot y^{-1} \neq y \cdot z^{-1} \ ,$$

and, for all pairwise distinct elements $x, y, z, w \in S$ it is

$$(ii) \qquad\qquad x \cdot y^{-1} \neq z \cdot w^{-1} \ .$$

Given a subset $W \subseteq G$, we form a hypergraph $\mathcal{G} = (W, \mathcal{E}_3 \cup \mathcal{E}_4)$ by collecting triples and quadruples of W which violate conditions (i) or (ii). For pairwise distinct elements $x, y, z \in W$ let $\{x, y, z\} \in \mathcal{E}_3$ if and only if $x \cdot y^{-1} = y \cdot z^{-1}$. Moreover, for pairwise distinct elements $x, y, z, w \in W$ let $\{x, y, z, w\} \in \mathcal{E}_4$ if and only if $x \cdot y^{-1} = z \cdot w^{-1}$. Constructing the sets $\mathcal{E}_3, \mathcal{E}_4$ can be done in time $O(n^3)$. Again, we want to find a large independent set in $\mathcal{G}$.
Let $\mathcal{G}|\mathcal{E}_4 = (W, \mathcal{E}_4)$. It is easy to see, cf. [33], that

$$|\mathcal{E}_3| \leq c_3 \cdot n^2 \ , \quad |\mathcal{E}_4| \leq c_4 \cdot n^3 \ , \quad s_{2,3}(\mathcal{G}|\mathcal{E}_4) \leq c_{2,3} \cdot n^3 \ , \quad \text{and} \quad s_{2,2}(\mathcal{G}|\mathcal{E}_4) \leq c_{2,2} \cdot n^4 \ ,$$

and the sets of $(2,3)$-cycles can be constructed in time $O(s_{2,3}(\mathcal{G}|\mathcal{E}_4))$. Moreover, for every subhypergraph $\mathcal{G}' = (W', \mathcal{E}_4')$ of $\mathcal{G}|\mathcal{E}_4$ we have

$$s_{2,2}(\mathcal{G}') \leq c_{2,2}' \cdot |\mathcal{E}'| \cdot |W'| \ .$$

The situation is similar to that in the proof of Theorem 5.9. We first choose with $p = n^{-1/3+\epsilon}$, where $0 < \epsilon < 1/3$, in time $O(n^3)$ a small subhypergraph $\mathcal{G}_0$ of $\mathcal{G}$, where we control the number of vertices, edges, and $(2,3)$-cycles in $\mathcal{G}|\mathcal{E}_4$. In the resulting hypergraph $\mathcal{G}_0$ the number of $(2,2)$-cycles is only $O(p^4 \cdot |\mathcal{E}_4| \cdot p \cdot n) = o(n^3)$ for $\epsilon < 2/15$. Then, on this small subhypergraph we choose in time $o(n^3)$ with $p = n^{-1/3+\epsilon_1}$ again a small subhypergraph $\mathcal{G}_1$, where we control the number of vertices, edges, and, among the four-element edges, the $(2,3)$-cycles and now also the $(2,2)$-cycles. In the resulting hypergraph $\mathcal{G}_1$ we delete in time $o(n)$ one vertex from each three-element edge, each $(2,2)$-cycle and each $(2,3)$-cycle, and we obtain a 4-uniform hypergraph without any 2-cycles, for which we apply in time $o(n^3)$ Corollary 2.8 and we get a desired Sidon set. □

Definition 5.15 *Let $(G, \cdot)$ be a group. Let $C \subset G$ be a subset of G with $1 \notin C$ where C is invariant under taking inverses, i.e., with $C = C^{-1} = \{c^{-1} \mid c \in C\}$. The* Cayley graph $\Gamma(G, C)$ *has vertex set G and edgeset $\{\{g, h\} \mid g \cdot h^{-1} \in C \ \text{ and } \ g, h \in G\}$.*

Corollary 5.16 *Let $(G, \cdot)$ be a group. There exists a constant $c > 0$ such that the following holds. Let H be a graph on n vertices. Then, for every subset $W \subseteq G$ with $|W| \geq c \cdot n^3 / \ln n$ one can construct in time $O\left(n^9/(\ln n)^3\right)$ some Cayley subgraph of G, and find in it an induced subgraph which is isomorphic to H.*

The existence of such a subset W with $|W| \geq c \cdot n^3$ was shown first by Babai and Sós [12]. This was improved to $|W| \geq c \cdot n^3/\log n$ in [33].

Proof: The arguments follow [12]. Let $H = (V, E)$ be a graph on n vertices and let $W \subseteq G$ be a subset of G with $|W| \geq c \cdot n^3/\ln n$, where $c > 0$ is a large enough constant. By Theorem 5.14 we obtain in time

$$O\left(\left(\frac{n^3}{\ln n}\right)^3\right) = O\left(\frac{n^9}{(\ln n)^3}\right)$$

a subset $S \subseteq W$ with $|S| = n$, where S is a Sidon set of the second kind. Now, we identify S with the vertex set V of H. Set $C = \{s \cdot t^{-1} \mid \{s,t\} \in E\}$. Then $1 \notin C$ and $C = C^{-1}$. Moreover, the Cayley graph $\Gamma(S, C)$ is an induced copy of the graph H. $\square$

In connection with the study of random Turán numbers the existence of graphs with many edges and without cycles of small lengths was shown by Kohayakawa, Kreuter and Steger [29]. We mention without proof that along the lines discussed in this paper the following can be shown:

Theorem 5.17 *Let k be a fixed positive integer. Then, one can compute in time $O(n^{4k-3})$ a graph G on n vertices, which does not contain any cycle $C_3, C_4, \ldots, C_{2k}$, and the number of edges in G is at least*

$$\Omega\left(n^{1+\frac{1}{2k-1}} \cdot (\ln n)^{\frac{1}{2k-1}}\right) .$$

Although there are better constructions known, for example, Ramanujan graphs arising from the work of Lubotzky, Phillips and Sarnak [35], i.e., graphs with at least $\Omega\left(n^{1+\frac{2}{3k+3}}\right)$ edges which do not contain any cycle $C_3, C_4, \ldots, C_{2k}$, the method used in [29] is interesting to gain a logarithmic factor. Namely, seemingly against the intuition they delete edges in a complete graph K_n with probability $p = n^{-1+\frac{1}{2k-1}+\epsilon}$ where $\epsilon > 0$ which is above the usual choice $p_0 = \Theta(n^{-1+\frac{1}{2k-1}})$. Then they form a hypergraph with vertices being the edges of the resulting graph, and (hyper-)edges consisting of the cycles $C_3, C_4, \ldots, C_{2k}$ of the graph K_n. The resulting hypergraph has an independent set of size $\Omega(n^{1+\frac{1}{2k-1}} \cdot (\ln n)^{\frac{1}{2k-1}})$, i.e., the corresponding edges in the graph form no cycle $C_3, C_4, \ldots, C_{2k}$.

For algorithmic reasons, we consider the complete graph on n vertices. We form a hypergraph $\mathcal{G}$ with vertices being the edges of K_n. The edges in $\mathcal{G}$ are determined by the edges of cycles of length at most $2k$ in K_n. Thus, the hypergraph has $\binom{n}{2}$ vertices and $\Theta(n^i)$ edges of cardinality i, where $i = 3, 4, \ldots, 2k$. Then we determine all 2-cycles among the $2k$-element edges of G. The number of $(2,j)$-cycles, $j = 2, 3, \ldots, 2k-1$ among the $2k$-element edges of G is $\Theta(n^{4k-j-1})$. Hence, the assumptions of Corollary 2.7 are fulfilled. We get rid of those edges with cardinality less than $2k$ by taking a small subhypergraph of $\mathcal{G}$, i.e., choosing vertices of $\mathcal{G}$ with probability $p = n^{-\frac{2k-2}{2k-1}+\epsilon}$ for some $\epsilon > 0$ and we obtain a $2k$-uniform subhypergraph for which we apply Corollary 2.7. The running time of this algorithm is given by the number of $(2,2)$-cycles, i.e., is $O(n^{4k-3})$.

References

[1] M. Ajtai, P. Erdös, J. Komlós and E. Szemerédi, *On Turán's Theorem for Sparse Graphs,* Combinatorica 1, 1981, 313-317.

[2] M. Ajtai, J. Komlós, J. Pintz, J. Spencer and E. Szemerédi, *Extremal Uncrowded Hypergraphs,* Journal of Combinatorial Theory Ser. A 32, 1982, 321-335.

[3] M. Ajtai, J. Komlós and E. Szemerédi, *A Note on Ramsey Numbers,* Journal of Combinatorial Theory, Ser. A 29, 1980, 354-360.

[4] M. Ajtai, J. Komlós and E. Szemerédi, *A Dense Infinite Sidon Sequence,* European Journal of Combinatorics 2, 1981, 2-11.

[5] N. Alon, *A Parallel Algorithmic Version of the Local Lemma,* Random Structures & Algorithms 2, 1991, 367-378.

[6] N. Alon, L. Babai and A. Itai, *A Fast and Simple Randomized Parallel Algorithm for the Maximal Independent Set Problem,* Journal of Algorithms 7, 1986, 567-583.

[7] N. Alon and P. Erdös, *An Application of Graph Theory to Additive Number Theory,* European Journal of Combinatorics 6, 1980, 201-203.

[8] N. Alon, H. Lefmann and V. Rödl, *On an Anti-Ramsey Type Result,* Colloquia Mathematica Societatis János Bolyai, 60. Sets, Graphs and Numbers, 1991, 9-22.

[9] N. Alon and J. Spencer, *The Probabilistic Method,* Wiley & Sons, New York, 1992.

[10] S. Arora, C. Lund, R. Motwani, M. Sudan and M. Szegedy, *Proof Verification and Hardness of Approximation Problems,* Proc. 33rd IEEE Symposium on Foundations of Computer Science FOCS, 1992, 14-23.

[11] L. Babai, *An Anti-Ramsey Theorem,* Graphs and Combinatorics 1, 1985, 23-28.

[12] L. Babai and V. T. Sós, *Sidon Sets in Groups and Induced Subgraphs of Cayley Graphs,* Graphs and Combinatorics, European Journal of Combinatorics 6, 1985, 101-114.

[13] M. Bellare, O. Goldreich and M. Sudan, *Free Bits, PCPs and Non-Approximability - Towards Tight Results*, Proc. 36th IEEE Symposium on Foundations of Computing FOCS, 1995, 422-431.

[14] R. Boppana and M. Halldórsson, *Approximating Maximum Independent Sets by Excluding Subgraphs,* BIT 32, 1992, 180-196.

[15] S. Brandt, *Computing the Independence Number of Dense Triangle-free Graphs,* to appear, 1997.

[16] R. A. Duke, H. Lefmann and V. Rödl, *On Uncrowded Hypergraphs,* Random Structures & Algorithms 6, 1995, 209-212.

[17] P. Erdös, R. L. Graham, I. Rusza and H. Taylor, *Bounds for Arrays of Dots with Distinct Slopes or Lengths,* Combinatorica 12, 1992, 39-44.

[18] P. Erdös and R. Guy, *Distinct Distances between Lattice Points,* Elemente der Mathematik 25, 1970, 121-123.

[19] A. Fundia, *Derandomizing Chebychev's Inequality to find Independent Sets in Uncrowded Hypergraphs*, Random Structures & Algorithms 8, 1996, 131-147.

[20] M. Goldberg and T. Spencer, *An Efficient Parallel Algorithm that Finds Independent Sets of Guaranteed Size,* SIAM Journal of Discrete Mathematics 6, 1993, 443-459.

[21] S. W. Golomb, *Construction of Signals with Favourable Correlation Properties,* Surveys in Combinatorics, London Mathematical Society Lecture Note Series 166, 1991, 1-39.

[22] S. W. Golomb and H. Taylor, *Two-dimensional Synchronization Patterns for Minimum Ambiguity,* IEEE Transactions Information Theory IT-28, 1982, 600-604.

[23] M. Halldórsson and J. Radhakrishnan, *Greed is Good: Approximating Independent Sets in Sparse and Bounded-degree Graphs,* Proc. 26th ACM Symposium on the Theory of Computing STOC, 1994, 439-448, also in: Algorithmica 18, 1997, 145-163.

[24] M. Halldórsson and J. Radhakrishnan, *Improved Approximations of Independent Sets in Bounded-degree Graphs via Subgraph Removal*, Nordic Journal on Computing 1, 1994, 475-492.

[25] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers,* Oxford University Press, 1979.

[26] J. Håstad, *Clique is Hard to Approximate Within* $n^{1-\epsilon}$*,* Proc. 37th IEEE Symposium on Foundations of Computer Science FOCS, 1996, 627-636.

[27] T. Hofmeister and H. Lefmann, *Derandomization for Sparse Approximations and Independent Sets*, Proc. Mathematical Foundations of Computer Science MFCS'95, LNCS 969, eds. J. Wiedermann and P. Hájek, Springer, 1995, 201–210.

[28] P. Kelsen, *On the Parallel Complexity of Computing a Maximal Independent Set in a Hypergraph,* Proc. 24th ACM Symposium on the Theory of Computing STOC, 1992, 339-350.

[29] Y. Kohayakawa, B. Kreuter and A. Steger, *An Extremal Problem for Random Graphs and the Number of Graphs with Large Even-Girth,* preprint, 1995.

[30] J. Komlós, J. Pintz and E. Szemerédi, *A Lower Bound for Heilbronn's Problem,* Journal of the London Mathematical Society 25, 1982, 13-24.

[31] G. Kortsarz and D. Peleg, *On Choosing a Dense Subgraph (Extended Abstract),* Proc. 34th IEEE Symposium on Foundations of Computer Science FOCS, 1993, 692-701.

[32] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen,* Teubner Verlag, Leipzig, 1909.

[33] H. Lefmann, V. Rödl and B. Wysocka, *Multicolored Subsets in Colored Hypergraphs,* Journal of Combinatorial Theory Ser. A 74, 1996, 209-248.

[34] H. Lefmann and T. Thiele, *Point Sets with Distinct Distances,* Combinatorica 15, 1995, 379-408.

[35] A. Lubotzky, R. Phillips and P. Sarnak, *Ramanujan Graphs,* Combinatorica 8, 1988, 261-277.

[36] K. T. Phelps and V. Rödl, *Steiner Triple Systems with Minimum Independence Number,* Ars Combinatorica 21, 1986, 167-172.

[37] P. Pudlák and V. Rödl, *Modified Ranks of Tensors and the Size of Circuits*, Proc. 25th ACM Symposium on the Theory of Computing STOC, 1993, 523-531.

[38] P. Raghavan, *Probabilistic Construction of Deterministic Algorithms: Approximating Packing Integer Programs,* Journal of Computer and System Sciences 37, 1988, 130-143.

[39] S. Ramanujan, *Collected Papers,* Chelsea Publishing Company, 1962, 133-135.

[40] V. Rödl and E. Siňajová, *Note on Independent Sets in Steiner Systems,* Random Structures & Algorithms 5, 1994, 183-190.

[41] J. Shearer, *A Note on the Independence Number of Triangle-free Graphs,* Discrete Mathematics 46, 1983, 83-87.

[42] J. Shearer, *A Note on the Independence Number of Triangle-free Graphs, II,* Journal of Combinatorial Theory 53, 1991, 300-307.

[43] J. Singer, *A Theorem in Finite Projective Geometry and some Applications to Number Theory,* Transactions of the American Mathematical Society 43, 1938, 377-385.

[44] J. Spencer, *Turán's Theorem for k-Graphs,* Discrete Mathematics 2, 1972, 183-186.

[45] J. Spencer, *Uncrowded Hypergraphs, in:* Mathematics of Ramsey Theory, eds. J. Nešetřil and V. Rödl, Springer, 1990, 253-262.

[46] T. Thiele, *Point Sets with Distinct Slopes or Lengths,* preprint, FU-Berlin, 1993.