Combinatorics, Probability and Computing (2025), 1–4
doi:10.1017/S0963548325000112

ARTICLE

Short proof of the hypergraph container theorem

Rajko Nenadov[1] and Huy Tuan Pham[2]

[1]School of Computer Science, University of Auckland, Auckland, New Zealand and [2]School of Mathematics, Institute for Advanced Study, Princeton, NJ, USA
Corresponding author: Huy Tuan Pham; Email: htpham@caltech.edu

(Received 11 November 2024; revised 17 February 2025; accepted 17 March 2025)

Abstract
We present a short and simple proof of the celebrated hypergraph container theorem of Balogh–Morris–Samotij and Saxton–Thomason. On a high level, our argument utilises the idea of iteratively taking vertices of largest degree from an independent set and constructing a hypergraph of lower uniformity which preserves independent sets and inherits edge distribution. The original algorithms for constructing containers also remove in each step vertices of high degree, which are not in the independent set. Our modified algorithm postpones this until the end, which surprisingly results in a significantly simplified analysis.

Keywords: Hypergraph containers; independent sets

2020 Mathematics Subject Classification: Primary: 05C30, 05C65, 05C69; Secondary: 05C35

1. Introduction

The method of containers is a powerful technique in combinatorics used to produce a small number of clusters encompassing independent sets of a given hypergraph. While in some applications, one follows the idea of the method and the general principles for building such clusters, quite often one can apply off-the-shelf tools. The most such applicable tool has been developed independently by Balogh, Morris, and Samotij [1] and Saxton and Thomason [9], and it is this result that is commonly referred to as the *hypergraph container theorem*. For an introduction to the method, the hypergraph container theorem, and its many surprising applications, we refer the reader to the International Congress of Mathematicians (ICM) survey [2]. A number of different proofs and versions of this result have been obtained since [3–5, 7, 8, 10, 11]. We present a simple and short proof of a slight generalisation of the original theorem. Two other short proofs have been obtained very recently by Campos and Samotij [6].

Let V be a finite set. Given a subset $X \subseteq V$, let $\langle X \rangle = \{S \subseteq V : X \subseteq S\}$. We say a probability measure ν over 2^V is (p, K)-*uniformly spread* if for every non-empty $X \subseteq V$, we have $\nu(\langle X \rangle) \leqslant Kp^{|X|-1}/|V|$. *Uniform* signifies that the measure is fairly uniform from the point of view of elements of V. Throughout the paper, we use $V = V(\mathcal{H})$ and $N = |V|$, where $\mathcal{H}$ is a given hypergraph. If all edges in a hypergraph $\mathcal{H}$ have size at most ℓ, we say that $\mathcal{H}$ is an$(\leqslant \ell)$-*graph*.

Theorem 1.1. *For every $\ell \in \mathbb{N}$ and $K, \varepsilon > 0$, there exists $T > 0$ such that the following holds. Suppose $\mathcal{H}$ is an $(\leqslant \ell)$-graph, and let ν be (p, K)-uniformly spread measure over 2^V supported on $\mathcal{H}$, for some $p \in (0, 1]$. Then for every independent set $I \subseteq V(\mathcal{H})$, there exists $F \subseteq I$ and $C = C(F) \subseteq V$ such that $|F| \leqslant TNp$, $\nu(\mathcal{H}[C]) < \varepsilon$, and $I \subseteq C \cup F$.*

If ν is uniform on $\mathcal{H}$, we obtain the original hypergraph container theorems [1, 9] (that being said, one can also obtain a non-uniform statement from the original containers by taking hyperedges with multiplicity). More importantly, it will be evident in our proof that non-uniform measures naturally arise and streamline the argument, even if one is only interested in proving Theorem 1.1 for the uniform measure. While we are not aware of an application where a non-uniform measure is used, it is conceivable that proving a supersaturation result, often used in a combination with container-type theorems, might become easier (or at least more elegant) if dealing with non-uniform measures. Dependence of T on the uniformity is of order $O(2^{\ell^2})$, which is also along the lines of the original results. Near-optimal dependence was obtained by Balogh and Samotij [3] and Campos and Samotij [6].

2. Proof

Our proof bears resemblance to the proof from [1, 9]. On a high level, we choose F in Theorem 1.1 by greedily taking vertices from I with largest degree with respect to ν and construct a hypergraph of lower uniformity given by (parts) of hyperedges containing vertices from F. A common feature in many of the proofs utilising a similar idea is that one also keeps track of the vertices which are not in I but have larger degree than the last chosen vertex in F. The main novelty here is that we completely avoid this, unless we are in a case where the resulting hypergraph of lower uniformity is not sufficiently dense to proceed with the induction. In this case, we show that removing vertices of high degree immediately yields a desired container. It is worth noting that the proofs from [1, 9] also have a similar case distinction; however; the analysis in our cases turns out to be significantly simpler.

Theorem 1.1 follows by iterated application of the following lemma, known as the *hypergraph container* lemma.

Lemma 2.1. *For every $\ell \in \mathbb{N}$ and $K > 0$, there exists $\delta > 0$ such that the following holds. Suppose $\mathcal{H}$ is an $(\leqslant \ell)$-graph, and there exists a (p, K)-uniformly spread measure ν over 2^V supported on $\mathcal{H}$, for some $p \in (0, 1]$. Then for every independent set $I \subseteq V$, there exists $F \subseteq I$ and $C = C(F) \subseteq V$ such that $|F| \leqslant \ell Np$, $|C| \leqslant (1-\delta)N$, and $I \subseteq C \cup F$. Moreover, C can be unambiguously constructed from any set $\hat{F}$ such that $F \subseteq \hat{F} \subseteq I$.*

Proof. We prove the lemma by induction on ℓ. For $\ell = 1$, take $F = \varnothing$ and $C \subseteq V$ to be the set of all vertices $v \in V$ with $\nu(v) = 0$. As there are at least N/K vertices with strictly positive measure, the lemma holds for $\delta = 1/K$. We now prove the lemma for $\ell \geqslant 2$. If $|I| < Np$, then we simply take $F = I$. Therefore, without loss of generality, we may assume $|I| \geqslant Np$.

Set $F = \varnothing \subseteq I$, $\mathcal{L} = \varnothing \subseteq 2^V$, and $\mathcal{D}, \mathcal{H}' = \varnothing \subseteq \mathcal{H}$. Here F denotes a "fingerprint"; $\mathcal{D}$ is the set of hyperedges we delete along the way due to some "heavy" sets (see (1), and $\mathcal{L}$ is the family of sets responsible for edges in $\mathcal{D}$ being deleted; the hypergraph $\mathcal{H}'$ consists of a subset of "nicely" distributed edges, which contain at least one vertex from F. Repeat the following for Np rounds: Take $v \in I \setminus F$ to be a largest vertex with respect to $\nu(\langle v\rangle \cap \mathcal{R})$, where $\mathcal{R} = \mathcal{H}[V \setminus F] \setminus \mathcal{D}$ (tie-breaking done in some canonical way, e.g. by agreeing on the ordering of V). Add v to F, set $\mathcal{H}' = \mathcal{H}' \cup (\langle v\rangle \cap \mathcal{R})$, and for each $X \in 2^V \setminus \mathcal{L}$ of size $|X| \leqslant \ell - 1$ such that

$$\nu(\langle X\rangle \cap \mathcal{H}') > Kp^{|X|}/N, \tag{1}$$

add X to $\mathcal{L}$ and set $\mathcal{D} = \mathcal{D} \cup (\langle X\rangle \cap \mathcal{R})$.

A few observations about the process. First, as ν is (p, K)-uniformly spread the value $\nu(\langle X\rangle \cap \mathcal{H}')$ increases by at most $\nu(\langle X \cup \{v\}\rangle) \leqslant Kp^{|X|}/N$ after adding a vertex $v \notin X$ to F. Once a subset X satisfies (1), no more hyperedges that contain X are added to $\mathcal{H}'$; thus at the end of the process; we have

$$\nu(\langle X\rangle \cap \mathcal{H}') \leqslant 2Kp^{|X|}/N \tag{2}$$

https://doi.org/10.1017/S0963548325000112 Published online by Cambridge University Press

for every $X \subseteq V \setminus F$ of size $|X| \leqslant \ell - 1$. Second, given a set $\hat{F}$ such that $F \subseteq \hat{F} \subseteq I$, we can reconstruct F from $\hat{F}$ together with the order in which the vertices were added; thus; we can also reconstruct $\mathcal{H}'$ and $\mathcal{R}$.

We next derive several useful lower bounds on $\nu(\mathcal{H}')$. First, we show that if $\nu(\mathcal{D})$ is large, then $\nu(\mathcal{H}')$ is also large. In particular, the following holds:

$$\nu(\mathcal{H}') \geqslant 2^{-\ell} p\nu(\mathcal{D}). \tag{3}$$

On the one hand, for each $e \in \mathcal{D}$, there exists $X \in \mathcal{L}$ such that $e \in \langle X \rangle$, thus $\sum_{X \in \mathcal{L}} \nu(\langle X \rangle) \geqslant \nu(\mathcal{D})$. On the other hand, we have by (1) that

$$\sum_{X \in \mathcal{L}} \nu(\langle X \rangle \cap \mathcal{H}') > \sum_{X \in \mathcal{L}} Kp^{|X|}/N \geqslant p \sum_{X \in \mathcal{L}} \nu(\langle X \rangle).$$

Here in the last inequality, we use that ν is (p, K)-uniformly spread. Furthermore, each edge e in $\mathcal{H}'$ may contribute to at most 2^{ℓ} terms $\nu(\langle X \rangle \cap \mathcal{H}')$. Hence,

$$\nu(\mathcal{H}') \geqslant 2^{-\ell} \sum_{X \in \mathcal{L}} \nu(\langle X \rangle \cap \mathcal{H}') > 2^{-\ell} p\nu(\mathcal{D}),$$

as claimed in (3).

Next, we show that

$$\nu(\mathcal{H}') \geqslant (Np) \max_{v \in I \setminus F} \nu(\langle v \rangle \cap \mathcal{R}). \tag{4}$$

Let $\mathcal{R}_i$ denote the hypergraph $\mathcal{R}$ at the moment when the i-th vertex v_i was added to F (thus $\mathcal{R} = \mathcal{R}_{|F|}$). We observe that since $\mathcal{R}$ is non-increasing and by our choice of v in each step,

$$\nu(\mathcal{H}') \geqslant \sum_{i=1}^{|F|} \nu(\langle v_i \rangle \cap \mathcal{R}_i) \geqslant \sum_{i=1}^{|F|} \max_{v \in I \setminus F} \nu(\langle v \rangle \cap \mathcal{R}_{|F|}),$$

yielding (4).

Let $\alpha = 2^{-\ell-2}$. We now distinguish two cases, where if $\nu(\mathcal{H}')$ is large, then we can apply the inductive hypothesis to an appropriate $(\leqslant \ell - 1)$-graph, and otherwise, we can immediately find a small container C for which $I \setminus F \subseteq C$.

Case 1: $\nu(\mathcal{H}') \geqslant \alpha p$. Let $\mathcal{H}''$ denote the $(\leqslant \ell - 1)$-graph consisting of sets X such that $X = H' \setminus F$ for some $H' \in \mathcal{H}'$. Set ν' to be the probability measure over $2^{V \setminus F}$ given by

$$\nu'(X) \propto \begin{cases} \nu((X \cup 2^F) \cap \mathcal{H}'), & \text{if } X \in \mathcal{H}'', \\ 0, & \text{otherwise}, \end{cases}$$

where $X \cup 2^F = \{X \cup Y \colon Y \in 2^F\}$. From (2) and $\nu(\mathcal{H}') \geqslant \alpha p$, we conclude that ν' is $(p, 2K\alpha^{-1})$-uniformly spread. Also observe that I is an independent set in $\mathcal{H}''$; thus, by the induction hypothesis, there exists $F' \subseteq V$ of size $|F'| \leqslant (\ell - 1)Np$ and $C = C(F')$ such that $|C| \leqslant (1 - \delta)N$ and $I \subseteq C \cup F'$. Note that we can reconstruct C from $F := F \cup F'$.

Case 2: $\nu(\mathcal{H}') < \alpha p$. By (3), we have $\nu(\mathcal{D}) < 1/4$ and hence $\nu(\mathcal{R}) \geqslant \nu(\mathcal{H}) - \nu(\mathcal{H}') - \nu(\mathcal{D}) > 1/2$. By (4), for every $v \in I \setminus F$, we have

$$\nu(\langle v \rangle \cap \mathcal{R}) \leqslant \alpha/N. \tag{5}$$

Let now $C \subseteq V \setminus F$ denote the set of all $v \in V \setminus F$ such that $\nu(\langle v \rangle \cap \mathcal{R}) \leqslant \alpha/N$. By (5), we have $I \setminus F \subseteq C$. Furthermore,

$$\nu(\mathcal{R}) \leqslant \sum_{v \in C} \nu(\langle v \rangle \cap \mathcal{R}) + \sum_{w \in V \setminus (F \cup C)} \nu(\langle w \rangle \cap \mathcal{R}) < \alpha + (N - |C|) \cdot K/N.$$

Hence, $|C| < N - (\nu(\mathcal{R}) - \alpha)N/K < (1 - \delta)N$ for $\delta = 1/(4K)$. This concludes the construction of the desired F and C. $\square$

https://doi.org/10.1017/S0963548325000112 Published online by Cambridge University Press

For the sake of completeness, we derive Theorem 1.1 from Lemma 2.1.

Proof of Theorem 1.1. Let $\delta > 0$ be as given by Lemma 2.1 for ℓ and K/ε (as K). We prove the theorem for $T = \ell \log(K\varepsilon^{-1})/\log(1+\delta)$.

We find a *fingerprint* F and a *container* C as follows. Set $F = \varnothing$ and $C = V$, and as long as $\nu(\mathcal{H}[C]) \geqslant \varepsilon$, do the following: Let F' and C' be as given by Lemma 2.1 applied with ν' being a probability measure over 2^C given by $\nu'(X) \propto \nu(X)$ if $X \in \mathcal{H}[C]$, and $\nu'(X) = 0$ otherwise. Set $F := F \cup F'$ and $C := C'$, and proceed to the next iteration.

If $\nu(\mathcal{H}[C]) \geqslant \varepsilon$, then for non-empty $X \subseteq C$,

$$\nu'(\langle X\rangle) \leqslant \frac{\nu(\langle X\rangle)}{\nu(\mathcal{H}[C])} \leqslant \frac{Kp^{|X|-1}/N}{\varepsilon} \leqslant \frac{K}{\varepsilon}p^{|X|-1}/|C|,$$

and hence ν' is $(p, K/\varepsilon)$-uniformly spread each time we apply Lemma 2.1. Furthermore, if $\nu(\mathcal{H}[C]) \geqslant \varepsilon$, then $|C| \geqslant \varepsilon N/K$. In each iteration, the set C shrinks by a factor of $1-\delta$; thus, we are done after at most $\log(K\varepsilon^{-1})/\log(1+\delta)$ iterations. The set F grows by at most ℓNp in each iteration, which gives an upper bound of TNp on its final size for the above choice of $T = T(K, \varepsilon)$. Due to the last property in Lemma 2.1, the final set C can be unambiguously constructed from F. $\square$

Acknowledgements

Ideas used in this paper were developed while the first author was visiting Stanford University in November 2023. The first author thanks Jacob Fox for his hospitality. We thank Jacob Fox for helpful comments and Wojciech Samotij and the anonymous referee for pointing out a subtle issue in an earlier version.

Research supported by the New Zealand Marsden Fund.

Research supported by a Clay Research Fellowship and a Stanford Science Fellowship.

References

[1] Balogh, J., Morris, R. and Samotij, W. (2015) Independent sets in hypergraphs. *J. Am. Math. Soc.* **28**(3) 669–709.

[2] Balogh, J., Morris, R. and Samotij, W. (2018) The method of hypergraph containers. In Proceedings of the International Congress of Mathematicians 2018, ICM 2018, Rio de Janeiro, Brazil, August 1–9, 2018, Vol. IV Invited lectures, World Scientific, Sociedade Brasileira de Matemática (SBM), Brazil, Hackensack, NJ-Rio de Janeiro, pp. 3059–3092.

[3] Balogh, J. and Samotij, W. (2020) An efficient container lemma. *Discrete Anal.* **2020** 56, Id/No 17.

[4] Bernshteyn, A., Delcourt, M., Towsner, H. and Tserunyan, A. (2019) A short nonalgorithmic proof of the containers theorem for hypergraphs. *Proc. Am. Math. Soc.* **147**(4) 1739–1749.

[5] Bucić, M., Fox, J. and Pham, H. T. (2024) Equivalence between Erdős-Hajnal and polynomial Rödl and Nikiforov conjectures, arXiv: 2403.08303.

[6] Campos, M. and Samotij, W. (2024) Towards an optimal hypergraph container lemma, arXiv: 2408.06617.

[7] Morris, R., Samotij, W. and Saxton, D. (2024) An asymmetric container lemma and the structure of graphs with no induced 4-cycle. *J. Eur. Math. Soc.* **26**(5) 1655–1711.

[8] Nenadov, R. (2024) Probabilistic hypergraph containers. *Israel J. Math.* **261**(2) 879–897.

[9] Saxton, D. and Thomason, A. (2015) Hypergraph containers. *Invent. Math.* **201**(3) 925–992.

[10] Saxton, D. and Thomason, A. (2016) Online containers for hypergraphs, with applications to linear equations. *J. Comb. Theory, Ser. B.* **121** 248–283.

[11] Saxton, D. and Thomason, A. (2016) Simple containers for simple hypergraphs. *Comb. Probab. Comput.* **25**(3) 448–459.

Cite this article: Nenadov R and Pham HT (2025). Short proof of the hypergraph container theorem. *Combinatorics, Probability and Computing*. https://doi.org/10.1017/S0963548325000112