1 Lecture 12: Lower Bounds for ε-Nets

Let $\mathcal{F}$ be a family of sets on a ground set U of size N. We say that $A \subseteq U$ is an *ε-net* if for every $F \in \mathcal{F}$ such that $|F| \geqslant \varepsilon N$, we have $F \cap A \neq \emptyset$. Given $\mathcal{F}$ and ε, it is natural to ask how small an ε-net can be. One cannot do better than $(1-\varepsilon)N$ in general, as this is necessary to cover all subsets of size εN. We can give better answers for families $\mathcal{F}$ that are simple in some sense; we formalize this with the notion of VC dimension. We say that a set $S \subseteq U$ is *shattered* by $\mathcal{F}$ if $|\{F \cap S : F \in \mathcal{F}\}| = 2^{|S|}$. We then define the *VC dimension of* $\mathcal{F}$ as $\mathrm{VC}(\mathcal{F}) = \max\{|S| : S \text{ shattered by } \mathcal{F}\}$.

As an example, consider a set P of $N \geqslant 3$ points in a plane. Take $\mathcal{F} = \{T \subseteq P : \exists \text{ a line } \ell \text{ such that } \ell \cap P = T\}$. It is easy to shatter a set of size 2, so $\mathrm{VC}(\mathcal{F}) \geqslant 2$. To see that $\mathrm{VC}(\mathcal{F}) \leqslant 2$, let $S \subseteq P$ be a set of size 3. If the points of S are not collinear, then there is no line that intersects all three points, so S cannot be shattered. Otherwise, denote the points of S by a, b, and c, where b lies on the segment between a and c. There is no line whose intersection with S is $\{a, c\}$ (any such line must also intersect b), so S cannot be shattered. The same argument applies to sets of size larger than 3, so $\mathrm{VC}(\mathcal{F}) = 2$.

Haussler and Welzl [22] showed that if $\mathrm{VC}(\mathcal{F}) = d$, then there is an ε-net for $\mathcal{F}$ of size $O\left(\frac{d}{\varepsilon}\log\frac{d}{\varepsilon}\right)$. This gives a net of size $O\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\right)$ for our example of lines in the plane. If we take our point set to be a disjoint union of $1/\varepsilon$ parallel lines, each with εN points, then every ε-net must be of size at least $1/\varepsilon$, because any ε-net has to contain at least one point from each of the parallel lines. In [33], Matoušek-Seidel-Welzl conjectured that "natural" geometric families $\mathcal{F}$ of VC dimension d admit an ε-net of size $O(d/\varepsilon)$. Alon [1] recently refuted this conjecture in the case of points and lines in the plane, using the density Hales-Jewett theorem to give a construction of a point set such that the smallest ε-net has size $\Omega\left(\frac{1}{\varepsilon}\log^*\frac{1}{\varepsilon}\right)$, where

$$\log^* n := \max\{i \in \mathbb{N} : \underbrace{\log \cdots \log}_{i \text{ times}} n > 1\}.$$

In particular, this bound is super-linear in $1/\varepsilon$.

In this lecture, our main goal will be to prove the following theorem of Balogh-Solymosi [12], which improves on the lower bound of Alon.

Theorem 1.1 ([12])**.** *Let $\varepsilon^* > 0$ be an arbitrary small constant. Then there exists a point system S in the plane, such that if $T \subseteq S$ intersects each line containing at least $\varepsilon^*|S|$ points, i.e., T is an ε^*-net, then*

$$|T| \geqslant \frac{1}{2\varepsilon^*} \log^{1/3}\left(\frac{1}{\varepsilon^*}\right) \log\log\left(\frac{1}{\varepsilon^*}\right)^{-1}.$$

It is conjectured that the true lower bound on the size of an ε-net for this setting should be $\Omega\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\right)$. The current state of the art is a lower bound of $\Omega\left(\frac{1}{\varepsilon}\log^{\frac{1}{2}-o(1)}\frac{1}{\varepsilon}\right)$.

1.1 Proof of Theorem 1.1

First, we recall the hypergraph container theorem for r-uniform hypergraphs. Let $\mathcal{H}$ be an r-uniform hypergraph with average degree d. For every $S \subseteq V(\mathcal{H})$, its co-degree, denoted by $d(S)$, is the number of edges in $\mathcal{H}$ containing S, i.e.,

$$d(S) = |\{e \in E(\mathcal{H}) : S \subseteq e\}|.$$

For every $j \in [r]$, denote by Δ_j the j-th maximum co-degree, i.e.,

$$\Delta_j = \max\{d(S) : S \subseteq V(\mathcal{H}), |S| = j\}.$$

For $\tau \in (0, 1)$, define

$$\Delta(\mathcal{H}, \tau) = 2^{\binom{r}{2}-1} \sum_{j=2}^{r} \frac{\Delta_j}{d\tau^{j-1} 2^{\binom{j-1}{2}}}.$$

We use the following version of the Saxton-Thomason [38] hypergraph container theorem.

Theorem 1.2 ([38, Corollary 3.6])**.** *Let $\mathcal{H}$ be an r-uniform hypergraph on vertex set $[N]$. Let $0 < \varepsilon, \tau < 1/2$. Suppose that $\tau < 1/(200 \cdot r \cdot r!^2)$ and $\Delta(\mathcal{H}, \tau) \leqslant \varepsilon/(12r!)$. Then there exists $c = c(r) \leqslant 1000 \cdot r \cdot r!^3$ and a collection of vertex susbset $\mathcal{C}$ such that*

1. *every independent set in $\mathcal{H}$ is a subset of some $A \in \mathcal{C}$;*

2. *for every $A \in \mathcal{C}$, $e(\mathcal{H}[A]) \leqslant \varepsilon \cdot e(\mathcal{H})$;*

3. *$\log|\mathcal{C}| \leqslant cN\tau \cdot \log(1/\varepsilon) \cdot \log(1/\tau)$.*

We now prove Theorem 1.1.

Proof of Theorem 1.1. We will take $V(\mathcal{H}) = [n]^r$ and $E(\mathcal{H})$ to be the set of collinear r-tuples, for r a parameter to be chosen later which satisfies $\log\log n \ll r \leqslant 0.01 \log^{1/2} n / \log\log n$. We then take a p-random subset of $\mathcal{H}$, i.e., we choose each vertex independently with probability p. From here, we obtain our final point set by removing one point from each collinear $(r+1)$-tuple in the p-random subset, resulting in a set with no collinear $(r+1)$-tuples. We then project this set onto the plane such that the collinearity of the points is maintained.

For this to go through, we want

$$\mathbb{E}[\#\text{ points}] = pn^r \gg p^{r+1}\frac{(r+1)\cdot 2^{2r+1}}{r!} n^{2r}\log n = \mathbb{E}[\#\text{ collinear }(r+1)\text{-tuples}],$$

which will be satisfied by taking $p = r/20n$.

We begin by setting parameters and computing some bounds arising from the hypergraph container theorem, Theorem 1.2. We will set $\varepsilon = 10^{-r^2}$ throughout.

Since there are at most n points in a line, we have for $2 \leqslant i < r$

$$\Delta_i = \binom{n-i}{r-i},$$

and $\Delta_r = 1$. We can bound (for details, see [12, Claim 4.1(i)] with $r = k$) the number of edges $e(\mathcal{H})$ and the average degree $d = d(\mathcal{H})$ as

$$\frac{n^{2r}}{r^{2r}} \leqslant e(\mathcal{H}) \leqslant \frac{2^{2r}}{(r-1)!}n^{2r} \quad \text{and} \quad \frac{n^r}{r^{2r-1}} \leqslant d = d(\mathcal{H}) \leqslant \frac{r^2\cdot 2^{2r}}{r!}n^r.$$

Taking $\tau = n^{-1-0.9/(r-1)}$, we can bound $\Delta(\mathcal{H},\tau)$ as

$$\Delta(\mathcal{H},\tau) \leqslant 2^{\binom{r}{2}-1}\sum_{j=2}^{r}\frac{\binom{n-j}{r-j}\cdot r^{2r-1}}{n^r\tau^{j-1}2^{\binom{j-1}{2}}} \leqslant \frac{\varepsilon}{12r!}.$$

Note also that since $r \leqslant \log^{1/2} n$, we have

$$\tau < \frac{1}{200\cdot r\cdot r!^2}.$$

We also need a supersaturation result for this hypergraph, which we take from Balogh-Solymosi [12]. Recall that $\varepsilon = 10^{-r^2}$.

Lemma 1.3 ([12, Lemma 6.1])**.** *Let $A \subseteq V(\mathcal{H})$ with $e(\mathcal{H}[A]) \leqslant \varepsilon\cdot e(\mathcal{H})$. Then $|A| \leqslant n^r/10$.*

With the above bounds, we can apply the hypergraph container theorem, Theorem 1.2, to obtain a set of containers $\mathcal{C}$ of size

$$|\mathcal{C}| \leqslant 2^{1000\cdot r\cdot r!^3 v(\mathcal{H})\tau\cdot\log(1/\varepsilon)\cdot\log(1/\tau)} \leqslant 2^{n^{-1-0.8/r}\cdot v(\mathcal{H})}.$$

Since every $C \in \mathcal{C}$ contains at most $\varepsilon \cdot e(\mathcal{H})$ edges, Lemma 1.3 implies that $|C| \leqslant 0.1 \cdot v(\mathcal{H})$ for all $C \in \mathcal{C}$. From here, we can upper bound the number of independent sets of size $m = \frac{9}{20} pv(\mathcal{H})$ in the p-random subset of $\mathcal{H}$ by counting the number of possible m-sets in each container and multiplying by the bound on the number of containers. Using the standard estimate $\binom{n}{k} \leqslant \left(\frac{en}{k}\right)^k$, we get that the number of independent sets of size m is bounded by

$$2^{n^{-1-0.8/r} \cdot v(\mathcal{H})} \cdot \binom{0.1 \cdot v(\mathcal{H})}{m} \cdot p^m \leqslant 2^{n^{-1-0.8/r} \cdot v(\mathcal{H})} (0.61)^m < 2^{(n^{-1-0.8/r} - p/5) \cdot v(\mathcal{H})} = o(1).$$

Denote by S the point set resulting from taking a p-random subset of $V(\mathcal{H})$, removing $o(pv(\mathcal{H}))$ points to ensure no $r+1$ points are in a line, and projecting this to $\mathbb{R}^2$. Standard concentration results imply that $|S| = (1 + o(1))pv(\mathcal{H})$ with high probability. As $\frac{9}{20} < \frac{1}{2}$, the previous bound implies that with high probability there are no independent sets in the p-random subset of $\mathcal{H}$ of size $|S|/2$. Note that if $T \subset S$ is an ε^*-net, then since T intersects every line containing r points in S and no line contains $r+1$ points of S, $S - T$ is an independent set. Then $|S| - |T| \leqslant |S|/2$, so $|T| \geqslant |S|/2$. Finally, we choose parameters by taking ε^* to be a small positive constant and setting

$$r := \log^{1/3}\left(\frac{1}{\varepsilon^*}\right) \log\log\left(\frac{1}{\varepsilon^*}\right)^{-1} \quad \text{and} \quad |S| = \frac{r}{\varepsilon^*}.$$

This gives us

$$|T| \geqslant \frac{|S|}{2} \geqslant \frac{1}{2\varepsilon^*} \log^{1/3}\left(\frac{1}{\varepsilon^*}\right) \log\log\left(\frac{1}{\varepsilon^*}\right)^{-1}. \qquad \square$$

Note that, in principle, one could compute the corresponding value of n in terms of the other parameters, but the existence of such n suffices. We remark here that

$$\varepsilon^* = 2^{-(1+o(1))r^3}, \quad n = 2^{(1+o(1))r^2}, \quad p = 2^{-(1+o(1))r^2}.$$

1.2 Weak ε-Nets

In geometric settings, the notion of an ε-net can be relaxed. Let P be a set of N points in the plane and L a set of lines. We say that $S \subset \mathbb{R}^2$ is a *weak-ε-net* if for every $\ell \in L$ such that $|\ell \cap P| \geqslant \varepsilon N$, we have $\ell \cap S \neq \emptyset$. The word *weak* here refers to the fact that S need not be a subset of P.

There are instances where weak-ε-nets can be much smaller than ε-nets. For example, take a collection of $1/\varepsilon$ distinct lines in the plane, all intersecting at a single point (say, the origin). Consider a point set obtained by taking εN points from each line, excluding the point of mutual intersection. Then any ε-net for this point set requires size $1/\varepsilon$, but the point at which all the lines intersect is a weak-ε-net of size 1.

Since an ε-net is a weak-ε-net by definition, there are weak-ε-nets of size $O\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\right)$ for the setting of points and lines in the plane. Alon gave a lower bound for weak-ε-nets in the planar points-and-lines setting in [1]. In [12], Balogh-Solymosi use ideas similar to those of Theorem 1.1 to give the following lower bound for weak-ε-nets.

Theorem 1.4 ([12]). *Let ε^0 be an arbitrary small constant. Then there is a $0 < \varepsilon^* < \varepsilon^0$ such that the following holds. There exists a point system S in the plane, that if T intersects each line containing at least $\varepsilon^*|S|$ points, i.e., T is a weak ε^*-net, then*

$$|T| \geqslant \frac{1}{10\varepsilon^*}\left(\log\log\frac{1}{\varepsilon^*}\right)^{1/10}.$$

1.3 Relation to Helly's Theorem

Recall Helly's theorem from convex geometry.

Theorem 1.5 (Helly). *For all finite families of convex sets in $\mathbb{R}^d$, if every $d+1$ sets have non-empty intersection, then the whole family has non-empty intersection.*

We can generalize Helly's theorem as follows. For $p \geqslant q \geqslant d+1$, we say that a family of convex sets $\mathcal{F}$ in $\mathbb{R}^d$ has the (p,q)-property if among every collection of p sets in $\mathcal{F}$, some q of them have non-empty intersection. In 1957, Hadwiger and Debrunner [21] posed the following conjecture.

Conjecture 1.6 ([21]). *For every $p \geqslant q \geqslant d+1$, there exists $c = c_d(p,q)$ such that any family $\mathcal{F}$ which satisfies the (p,q)-property can be pierced by c points. That is, there exists a set $S \subset \mathbb{R}^d$ with $|S| = c$ such that for all $F \in \mathcal{F}$, we have $F \cap S \neq \emptyset$.*

Note that Helly's theorem immediately implies that $c_d(d+1,d+1) = 1$. It is not hard to see that $c_d(p,q) \geqslant p-q+1$. Indeed, the union of many sets overlapping at a point and $p-q$ disjoint sets shows that $p-q+1$ points are necessary. Hadwiger and Debrunner [21] showed a matching upper bound when $q \geqslant p/2+1$. Subsequently, Alon and Kleitman [2] proved that $c_d(p,q)$ is finite, giving an exponential bound. In recent work, Keller and Smorodinsky [25] gave a lower bound of $p^{1+\Omega(1/q)}$ on $c_2(p,q)$. A special case of their work is the following theorem, for which we sketch the proof. Further details can be found in [25].

Theorem 1.7 ([25]). $c_2(p,3) \geqslant p^{6/5-o(1)}$.

Proof sketch. Let S be a set of n points in the plane which contains no four points on a line. By a result of Balogh-Solymosi [12, Theorem 2.1] on the (3,4)-problem (the problem of finding the minimal size of the largest $S' \subseteq S$ such that S' is in general position, for any set S satisfying the above conditions), we can assume S has the additional property that for any $R \subseteq S$, if $|R| \geqslant n^{5/6+o(1)}$ then R contains 3 points on a line. Via point-line duality, we get a collection $\mathcal{F}$ of n lines such that no 4 lines have a common point and that every set of $n^{5/6+o(1)}$ lines contains three lines sharing a common point. Take $p = n^{5/6+o(1)}$. Then $\mathcal{F}$ satisfies the $(p,3)$-property. Moreover, every point intersects at most 3 lines, so at least $n/3 = p^{6/5-o(1)}$ points are needed to pierce $\mathcal{F}$. $\square$

References

[1] N. Alon, *A non-linear lower bound for planar epsilon-nets*, Proc. of the 51st IEEE FOCS (2010), 341–346. Also: *Discrete and Computatioanl Geometry*, **47** (2012), 235–244.

[2] N. Alon and D.J. Kleitman. *Piercing convex sets and the Hadwiger-Debrunner (p,q)-problem.* Advances in Mathematics, 96(1):103–112, 1992.

[3] N. Alon and J.H. Spencer, *T*he probabilistic method, third ed., Wiley-Intersecience in Discrete Math. and Optimization, John Wiley & Sons Inc., Hoboken, NJ, 2008, With an appendix on the life and work of Paul Erdős.

[4] J. Balogh, N. Bushaw, M. Collares, H. Liu, R. Morris and M. Sharifzadeh (2017). *The typical structure of graphs with no large cliques.* Combinatorica, 37(4), 617–632.

[5] J. Balogh, R. Morris and W. Samotij, *Independent sets in Hypergraphs*, J. Amer. Math. Soc. 28 (2015), no. 3, 669–709

[6] J. Balogh, R. Morris, and W. Samotij, *The Method of Hypergraph Containers*, 2018, survey for ICM 2018

[7] J. Balogh, R. Morris, W. Samotij and L. Warnke, 'The typical structure of sparse K_{r+1}-free graphs', *Transactions of the American Mathematical Society*, 368(9) (2013).

[8] J. Balogh, R. Mycroft, and A. Treglown, *A random version of Sperner's theorem*, Journal of Combinatorial Theory, Series A (2014) 128, 104–110.

[9] J. Balogh, H. Liu, S. Petříčková, and M. Sharifzadeh 'Typical structure of maximal triangle-free graphs', *Forum of Mathematics,* Sigma, 3, e20 (2015).

[10] J. Balogh, H. Liu and M. Sharifzadeh, 2016. *The number of subsets of integers with no k-term arithmetic progression.* International Mathematics Research Notices, 2017(20), pp.6168–6186.

[11] J. Balogh and S. Petříčková, 'The number of the maximal triangle-free graphs', *Bull. London Math. Soc.* 46 (5) (2014) 1003–1006.

[12] J. Balogh and J. Solymosi, *On the number of points in general position in the plane*, Discrete Analysis, 2018:16, 20pp.

[13] P. J. Cameron and P. Erdős. *On the number of sets of integers with various properties.* Number theory (Banff, AB, 1988), 61–79. (1990)

[14] D. Conlon and W. T. Gowers, *Combinatorial theorems in sparse random sets*, Ann. Math., Second Series, 184 no. 2 (2016), 367–454.

[15] P. Erdős, *Some recent results on extremal problems in graph theory.* Results, Theory of Graphs (Internat. Sympos., Rome, 1966), Gordon and Breach, New York; Dunod, Paris, 1967, pp. 117123 (English); pp. 124–130 (French).

[16] P. Erdős, D. J. Kleitman and B. L. Rothschild, 'Asymptotic enumeration of K_n-free graphs', *Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), Tomo II, Atti dei Convegni Lincei 17, Accad. Naz. Lincei* (1976) 19–27.

[17] P. Erdős and M. Simonovits, *Supersaturated Graphs and Hypergraphs*. Combinatorica 3 (1983), 183.

[18] D. Galvin, *Counting Colorings of a Regular Graph*, Graphs and Combinatorics (2015) 31, 629–638.

[19] D. Galvin and P. Tetali, *On weighted graph homomorphisms*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science (2004) 63, Graphs, Morphisms and Statistical Physics, 97–104.

[20] B. Green and T. Tao, *T*he primes contain arbitrarily long arithmetic progressions, Ann. Math., 167 no. 2 (2008) 481–547.

[21] H. Hadwiger and H. Debrunner. *Über eine variante zum Hellyschen satz.* Archiv der Mathematik, 8(4):309–313, 1957.

[22] D. Haussler and E. Welzl, *ε-nets and simplex range queries*, Discrete and Computational Geometry, (1987), 127–151.

[23] M. Hujter and Z. Tuza, 'The Number of Maximal Independent Sets in Triangle-Free Graphs', *SIAM J. Discrete Math.* 6(2) (1993) 284–288.

[24] J. Kahn, *An entropy approach to the hard-core model on bipartite graphs*, Combinatorics, Probability and Computing (2001) 10, 219–237.

[25] C. Keller and S. Smorodinsky, *A New Lower Bound on Hadwiger-Debrunner Numbers in the Plane*, arXiv:1809.06451.

[26] D. Kleitman, *A conjecture of Erdős-Katona on commmensurable pairs among subsets of an n-set*, Theory of Graphs, Proc. Colloq., Tihany (1966), 215–218.

[27] D. Kleitman, *On Dedekind's problem: the number of monotone boolean functions*, Proc. Amer. Math. Soc. (1969) 21, 677–682.

[28] D. J. Kleitman and K. J. Winston, *The asymptotic number of lattices*, Ann. Discrete Math. 6 (1980), 243–249, Combinatorial mathematics, optimal designs and their applications (Proc. Sympos. Combin. Math. and Optimal Design, Colorado State Univ., Fort Collins, Colo., 1978).

[29] D. J. Kleitman and K. J. Winston, *On the number of graphs without 4-cycles*, Discrete Math. 41 (1982), 167–172.

[30] Y. Kohayakawa, S. J. Lee, V. Rödl, W. Samotij, *The number of Sidon sets and the maximum size of Sidon sets contained in a sparse random set of integers.* Random Structures & Algorithms 46, no. 1 (2015): 1–25.

[31] J. Moon and L. Moser, 'On cliques in graphs', *Israel J. Math.*, (3) (1965) 23–28.

[32] W. Mantel, 'Problem 28', *Wiskundige Opgaven* 10 (1907) 60–61.

[33] J. Matoušek, R. Seidel and E. Welzl, *How to net a lot with little: Small ε-nets for disks and halfspaces*, In Proc. 6th Annu. ACM Sympos. Comput. Geom., 16–22 1990.

[34] D. Osthus, H.J. Prömel and A. Taraz, 'For which densities are random triangle-free graphs almost surely bipartite?' *Combinatorica* 23 (2003) 105–150.

[35] D. Osthus, *Maximum antichains in random sets of a finite set*, Journal of Combinatorial Theory, Series A (2000) 90, 36–46.

[36] I. Z. Ruzsa and E. Szemerédi, 'Triple systems with no six points carrying three triangles', *Combinatorics* (Keszthely, 1976), *Coll. Math. Soc. J. Bolyai* 18, Vol. II, (1978) 939–945.

[37] A. A. Sapozhenko, *The number of antichains in ranked partially ordered sets*, Diskret. Math. (1989) 1, 74–93.

[38] D. Saxton and A.G. Thomason, *Hypergraph Containers*. Inventiones Mathematicae, 201 (2015), 925–992.

[39] M. Schacht, *Extremal results for random discrete structures*, Ann. Math., Second Series, 184 no. 2 (2016), 333–65.

[40] E. Sperner, *Ein Satz über Untermengen einer endlichen Menge*, Mathematische Zeitschrift (1928) 27, 544–548.

[41] M. Simonovits, *A method for solving extremal problems in graph theory*, stability problems, Theory of Graphs (Proc. Colloq., Tihany, 1966), Academic Press, New York, 1968, pp. 279–319.

[42] M. Simonovits, 'Paul Erdős' influence on extremal graph theory', *The Mathematics of Paul Erdős* II (Springer-Verlag, Berlin, 1996) 148–192.

[43] E. Szemerédi, *O*n sets of integers containing no k elements in arithmetic progression, Acta Arith. 27 (1975), 199–234.

[44] P. Turán, *Eine Extremalaufgabe aus der Graphentheorie*, Mat. Fiz. Lapok 48 (1941), 436–452.

[45] P. Varnavides, *O*n certain sets of positive density, J. London Math. Soc. 34 (1959), 358–360.

[46] Y. Zhao, *The number of independent sets in a regular graph*, Combinatorics, Probability and Computing (2010) 19, 315–320.