HAL
open science

A Simple Proof of the Shallow Packing Lemma

Nabil Mustafa

To cite this version:

Nabil Mustafa. A Simple Proof of the Shallow Packing Lemma. Discrete and Computational Geometry, 2016, 55 (3), pp.739-743. 10.1007/s00454-016-9767-5 . hal-01345858

HAL Id: hal-01345858

https://hal.science/hal-01345858

Submitted on 16 Jul 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Simple Proof of the Shallow Packing Lemma

Nabil H. Mustafa*

Abstract

We show that the shallow packing lemma follows from a simple modification of the standard proof, due to Haussler and simplified by Chazelle, of the packing lemma.

1 Introduction

In 1995 Haussler [5] proved the following interesting theorem, improving upon an earlier result of Dudley [2].

Theorem A (Packing Lemma [6, Lemma 5.14])**.** *Let $(X, \mathcal{P})$ be a set-system on n elements, and with VC-dimension at most d. Let δ be an integer, $1 \leq \delta \leq n$, such that $|\Delta(R, S)| \geq \delta$ for every $R, S \in \mathcal{P}$, where $\Delta(R, S) = (R \setminus S) \cup (S \setminus R)$. Then $|\mathcal{P}| = O\big((n/\delta)^d\big)$.*

Haussler's proof is a beautiful application of the probabilistic method (in particular, conditional variance), and was simplified by Chazelle [1]. Recently much effort has been devoted to finding size-sensitive generalizations of this result. Given a set system $(X, \mathcal{P})$ and an integer $\delta > 0$, we say that $(X, \mathcal{P})$ is *δ-separated* if $|\Delta(R, S)| \geq \delta$ for every $R, S \in \mathcal{P}$. For any set $Y \subseteq X$, define the *projection* of $\mathcal{P}$ onto Y as the set system $\mathcal{P}|_Y = \big\{S \cap Y \mid S \in \mathcal{P}\big\}$. After a series of partial bounds [8, 4, 7], the following statement has been recently established in [3], via two different proofs (one building on Haussler's original proof while the other extends Chazelle's proof):

Theorem B (Shallow Packing Lemma)**.** *Let $(X, \mathcal{P})$ be a set-system on n elements, and let $d, d_1, k, \delta > 0$ be integers. Assume that $\mathcal{P}$ has VC-dimension at most d. Further, assume that for any set $Y \subseteq X$ the number of sets in $\mathcal{P}|_Y$ of size at most l is at most $f(|Y|, l) = O\big(|Y|^{d_1} l^{d-d_1}\big)$. If $\mathcal{P}$ is δ-separated and $|S| \leq k$ for all $S \in \mathcal{P}$, then $|\mathcal{P}| = O\big(n^{d_1} k^{d-d_1}/\delta^d\big)$.*

The objective of this paper is to prove that Theorem B (in fact, a generalization of it), with a simple trick, is a consequence of Haussler and Chazelle's Packing Lemma, which we first state below in a slightly more general form.

Theorem 1. *Let $(X, \mathcal{P})$ be a set-system on n elements. Let d, δ be two integers such that the VC-dimension of $\mathcal{P}$ is at most d, and $\mathcal{P}$ is δ-separated. Then*

$$|\mathcal{P}| \leq 2 \cdot \mathrm{E}\left[|\mathcal{P}|_{A'}|\right], \textit{where } A' \textit{ is a uniformly chosen random sample of size } \frac{4dn}{\delta} - 1.$$

*Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge, Equipe A3SI, ESIEE Paris. E-email: mustafan@esiee.fr. The work of Nabil H. Mustafa in this paper has been supported by the grant ANR SAGA (JCJC-14-CE25-0016-01).

Using it, Theorem B can be proven in a more general setting in terms of the so-called *shallow-cell complexity* of set systems. Given a function $\varphi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, a set system $(X, \mathcal{P})$ has shallow-cell complexity $\varphi(\cdot, \cdot)$ if for any $Y \subseteq X$, the number of subsets in $\mathcal{P}|_Y$ of size at most l is at most $|Y| \cdot \varphi\big(|Y|, l\big)$. Our main result is the following.

Theorem 2. *Let $(X, \mathcal{P})$ be a set-system, and $d, k, \delta > 0$ be integers. Assume that $|X| = n$, the VC-dimension of $\mathcal{P}$ is at most d, and $\mathcal{P}$ has shallow-cell complexity $\varphi(\cdot, \cdot)$. If $\mathcal{P}$ is δ-separated and $|S| \leq k$ for all $S \in \mathcal{P}$, then $|\mathcal{P}| = O\Big(\frac{n}{\delta} \cdot \varphi\big(\frac{4dn}{\delta}, \frac{12dk}{\delta}\big)\Big)$.*

Note that if for any set $Y \subseteq X$ the number of sets in $\mathcal{P}|_Y$ of size at most l is $O(|Y|^{d_1} l^{d-d_1})$, then $\mathcal{P}$ has shallow-cell complexity $\varphi(n, l) = O\big(n^{d_1-1} l^{d-d_1}\big)$, and so Theorem 2 implies Theorem B.

Organization. We prove the main theorem in Section 2. As the proof follows from a slight generalization of Haussler and Chazelle's proof of the packing lemma (as stated in Theorem 1), we present its proof in Section 3.

2 Proof of Theorem 2.

The proof in [6, Lemma 5.14] proves Theorem A as follows. By the primal shatter lemma [6], we have $|\mathcal{P}|_Y| = O((n/\delta)^d)$, and from Theorem 1 we can conclude that $m = |\mathcal{P}| = O((n/\delta)^d)$. Now we show that the proof of Theorem 2 is also a similar step away, by using instead the shallow-cell complexity of the set system.

Proof of Theorem 2. Let $A' \subseteq X$ be a uniform random sample of size $\frac{4dn}{\delta} - 1$. Define

$$\mathcal{P}_1 = \{S \in \mathcal{P} \text{ s.t. } |S \cap A'| > 3 \cdot 4dk/\delta\}.$$

Note that $\mathrm{E}[|S \cap A'|] \leq 4dk/\delta$ as $|S| \leq k$ for all $S \in \mathcal{P}$. By Markov's inequality, for any $S \in \mathcal{P}$,

$$\Pr[S \in \mathcal{P}_1] = \Pr[|S \cap A'| > 3 \cdot 4dk/\delta] \leq 1/3.$$

Thus

$$\begin{aligned} \mathrm{E}[|\mathcal{P}|_{A'}|] &\leq \mathrm{E}[|\mathcal{P}_1|] + \mathrm{E}[|(\mathcal{P} \setminus \mathcal{P}_1)|_{A'}|] \\ &\leq \sum_{S \in \mathcal{P}} \Pr[S \in \mathcal{P}_1] + |A'| \cdot \varphi\Big(|A'|, \frac{12dk}{\delta}\Big) \\ &\leq \frac{|\mathcal{P}|}{3} + \frac{4dn}{\delta} \varphi\Big(\frac{4dn}{\delta}, \frac{12dk}{\delta}\Big) \end{aligned}$$

where the projection size of $\mathcal{P} \setminus \mathcal{P}_1$ to A' is bounded by $\varphi(\cdot, \cdot)$. Now the bound follows from Theorem 1.

□

3 Proof of Theorem 1.

For the sake of completeness, we reproduce the proof of Haussler and Chazelle, noting a slight generalization of it as stated in Theorem 1. Alternatively, it can be found in the textbook [6, Lemma 5.14].

Given $\mathcal{P} = \{S_1, \ldots, S_m\}$ on a set X of n elements, we first define the *unit distance* graph $G_U(\mathcal{P}) = (\mathcal{P}, E_{\mathcal{P}})$. The vertex set of $G_U(\mathcal{P})$ is $\mathcal{P}$, and for any $S_i, S_j \in \mathcal{P}$, $\{S_i, S_j\} \in E_{\mathcal{P}}$ if and only if $|\Delta(S_i, S_j)| = 1$. We will need the following result on unit distance graphs of set systems.

Lemma 3 (Haussler [5]). *Given a set system $(X, \mathcal{P})$ with VC-dimension d, let $G_U(\mathcal{P}) = (\mathcal{P}, E_{\mathcal{P}})$ be its unit distance graph. Then $|E_{\mathcal{P}}| \leq d|\mathcal{P}|$.*

Proof of Theorem 1. Let $\mathcal{P} = \{S_1, \ldots, S_m\}$, where $m = |\mathcal{P}|$. Choose A to be a random sample of X of size $s = 4dn/\delta$, picked uniformly from all s-sized subsets of X. Let $G_U(\mathcal{P}|_A) = (\mathcal{P}|_A, E_{\mathcal{P}|_A})$ be the unit distance graph on $\mathcal{P}|_A$. For any set $S' \in \mathcal{P}|_A$, define

$$w(S') = \big|\{S \in \mathcal{P} \;\; s.t. \;\; S \cap A = S'\}\big|.$$

Define the weight of an edge $\{S'_i, S'_j\} \in E_{\mathcal{P}|_A}$ to be $w(\{S'_i, S'_j\}) = \min\{w(S'_i), w(S'_j)\}$. Let $W = \sum_{e \in E_{\mathcal{P}|_A}} w(e)$.

Claim 3.1. $W \leq 2d \cdot m$.

Proof. By Lemma 3, $|E_{\mathcal{P}|_A}| \leq d|\mathcal{P}|_A|$. Hence there exists a set $S' \in \mathcal{P}|_A$ with degree at most $2d$ in $G_U(\mathcal{P}|_A)$. Furthermore the weight of each edge incident to S' is at most $w(S')$, thus the total weight of edges incident to S' is at most $2d \cdot w(S')$. Remove S' from $G_U(\mathcal{P}|_A)$, and recursively bound the weight of edges in the remaining graph. Thus the total weight of edges is at most $2d \sum_{S'} w(S') = 2d \cdot m$. □

An alternate way of picking A is by first choosing randomly a set A' of $s-1$ elements of X, and then choosing the last element uniformly from $X \setminus A'$. Let W_1 be the random variable denoting the weight of the edges in $G_{\mathcal{P}|_A}$ for which the element in $A \setminus A'$ is the symmetric difference. By symmetry, we have

$$\mathrm{E}[W] = s \cdot \mathrm{E}[W_1] \tag{1}$$

To compute $\mathrm{E}[W_1]$, fix the set of first $s-1$ vertices. Now conditioned on this fixed choice of A', we show the following statement.

Claim 3.2. $\mathrm{E}\left[W_1 | A' = Y\right] \geq \frac{\delta}{n}\Big(m - |\mathcal{P}|_Y|\Big)$.

Proof. Consider a set $S' \in \mathcal{P}|_Y$, and let $\mathcal{P}_{S'}$ be the sets of $\mathcal{P}$ whose projection to Y is S'. Once the choice of the last element a has been made, S' will be split into two sets S'_1 and S'_2 in $\mathcal{P}|_A$, where $S'_1 = S'$ and $S'_2 = S' \cup \{a\}$. Similarly $\mathcal{P}_{S'}$ will be partitioned into $\mathcal{P}_{S'_1}$, consisting of the sets of $\mathcal{P}$ whose projection to A is S'_1 and $\mathcal{P}_{S'_2}$, consisting of the sets of $\mathcal{P}$ whose projection to A is $\mathcal{P}_{S'_2}$. Let $b_1 = |\mathcal{P}_{S'_1}|$ and $b_2 = |\mathcal{P}_{S'_2}|$. The weight of the edge in $G_U(\mathcal{P}|_A)$ between the two sets S'_1 and S'_2 is $\min\{b_1, b_2\}$, and it is shown in [6, Lemma 5.14] that

$$\mathrm{E}\left[\min\{b_1, b_2\}\right] \geq \frac{\delta}{n} \cdot \Big(|\mathcal{P}_{S'}| - 1\Big)$$

Summing up over all sets of $\mathcal{P}|_Y$,

$$\mathrm{E}\left[W_1 | A' = Y\right] \geq \sum_{S' \in \mathcal{P}|_Y} \frac{\delta}{n}\Big(|\mathcal{P}_{S'}| - 1\Big) = \frac{\delta}{n}\Big(m - |\mathcal{P}|_Y|\Big).$$

□

From Equation (1) and Claim 3.2, we get a lower-bound on $\mathrm{E}[W]$:

$$\begin{aligned}
\mathrm{E}[W] &= s \cdot \mathrm{E}[W_1] = s \cdot \sum_{\substack{Y \subseteq X \\ |Y| = s-1}} \mathrm{E}[W_1 | A' = Y] \cdot \Pr[A' = Y] \\
&\geq s \cdot \sum_{\substack{Y \subseteq X \\ |Y| = s-1}} \frac{\delta}{n}\Big(m - |\mathcal{P}|_Y|\Big) \cdot \Pr[A' = Y] \\
&\geq \frac{s\delta}{n}\left(m \sum_{\substack{Y \subseteq X \\ |Y| = s-1}} \Pr[A' = Y] - \sum_{\substack{Y \subseteq X \\ |Y| = s-1}} |\mathcal{P}|_Y| \cdot \Pr[A' = Y]\right) \\
&= 4dm - 4d\,\mathrm{E}[|\mathcal{P}|_{A'}|]
\end{aligned}$$

where the second inequality follows from Claim 3.2. Together with the upper-bound on W from Claim 3.1, we get

$$2dm \geq \mathrm{E}[W] \geq 4dm - 4d\,\mathrm{E}[|\mathcal{P}|_{A'}|], \text{ implying that } m \leq 2\,\mathrm{E}[|\mathcal{P}|_{A'}|] \text{ as desired.}$$

□

Discussion. Besides a shorter proof, we have shown that a generalization of the shallow packing lemma follows without any modification or addition to the proof of Haussler and Chazelle. The somewhat subtle key idea that was missed by earlier work [3, 4] is that it is fine if $\mathrm{E}[|\mathcal{P}|_{A'}|]$ is bounded in terms of $c|\mathcal{P}|$ for a small-enough constant c, as in any case it would be absorbed by the LHS of the equation in Theorem 1. This allows us to replace the complicated technical machinery developed in earlier work (iterative processes, Chernoff bounds for hypergeometric series, complicated probabilistic computations) by a mere Markov's inequality.

References

[1] Bernard Chazelle. A note on Haussler's packing lemma. 1992.

[2] Richard M. Dudley. Central limit theorems for empirical measures. *Ann. Probab.*, 6(6):899–1049, 1978.

[3] Kunal Dutta, Esther Ezra, and Arijit Ghosh. Two proofs for shallow packings. In *31st International Symposium on Computational Geometry (SoCG)*, pages 96–110, 2015.

[4] Esther Ezra. A size-sensitive discrepancy bound for set systems of bounded primal shatter dimension. In *Proc. of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1378–1388, 2014.

[5] David Haussler. Sphere packing numbers for subsets of the boolean n-cube with bounded Vapnik-Chervonenkis dimension. 69(2):217 – 232, 1995.

[6] Jiří Matoušek. *Geometric Discrepancy : An Illustrated Guide*. Algorithms and combinatorics. Springer, Berlin, New York, 1999.

[7] Nabil H. Mustafa and Saurabh Ray. Near-optimal generalisations of a theorem of Macbeath. In *31st International Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 578–589, 2014.

[8] Evangelia Pyrga and Saurabh Ray. New existence proofs for epsilon-nets. In *Symposium on Computational Geometry*, pages 199–207, 2008.