A little from [IJKW08]

<u>Inequalities</u> – Let $X_1,...,X_t$ be random variables taking values in the interval $[0,1]$, with expectation μ_i. Let $X=\sum_{i=1}^{t} X_i$ and $\mu=\sum_{i=1}^{t}\mu_i$ the expectation of X. Then for any $0<\gamma\le 1$ we have:

- *Chernoff-Hoeffding* – if $X_1,...,X_t$ are independent then
  $\Pr\left[|X-\mu|\ge\gamma\mu\right]\le 2\cdot e^{-\frac{\gamma^2\mu}{3}}$.

- *Chebyshev* – if $X_1,...,X_t$ are pairwise independent then
  $\Pr\left[|X-\mu|\ge\gamma\mu\right]\le\frac{1}{\gamma^2\mu}$
  כלומר, ההסת' שסכום איברים יהיה רחוק מהממוצע שלו ב $\gamma\mu$ קטן ממה שכתוב בימין.

<u>Bi-regular (bi-partite) graph</u> – The degree of all vertices in L is the same, and the degree of all in R is the same.

<u>Inclusion Graph</u> – T family of k-subsets, S family of s-subsets, s<k. Bi-partite graph that has an edge $(s\in S, t\in T)$ iff $s\subseteq t$.

<u>S-Graph</u> – For every $B\in$ T, the S-Graph $H(B,N(B))$, has each $x\in B$ on one side, connected to all the $A\in N(B)$, such that $(x,A)\in E \Leftrightarrow x\in A$.

כלומר, פריסה של קבוצה אחת מ T, שמצביעה לכל איבר על כל הקבוצות מ S (המוכלות ב T) שמכילות אותו.

<u>T-Graph</u> – Opposite graph to S-Graph. It is $G(U\setminus A, N(A))$, where every member x of $U\setminus A$ points to the $B\in N(A)$ such that $x\in B\setminus A$.

כלומר פריסה של כל המרחב מלבד קבוצה אחת מ S, שמצביעה לכל איבר על הקבוצות השכנות ל A שמכילות אותו. (מתקיים כי האיבר x שייך בעצם ל $B\setminus A$ מכיוון שמראש נלקח מ $U\setminus A$).

<u>Transitive Graph</u> –

קיימת פרמוטציה שמעבירה של אחד לכל אחד אחר בגדול. ואז מבנה הגרפים לעיל נשאר אותו דבר בלי תלות בבחירת A או B.

Samplers – Let $G=G(L,R)$ be a bi-partite, bi-regular graph. For $\lambda:[0,1]\to[0,1]$ we say that G is a $(\beta,\lambda(\beta))-sampler$ if for every function $F:L\to[0,1]$, where the average value μ is defined $\mu \triangleq \mathrm{Exp}_{x\in L}\left[F(x)\right]$

(כלומר ממוצע על כל הערכים ש F מקבלת)

There are at most $\lambda(\mu)\cdot|R|$ vertices $r\in R$ where

$$\left|\mathrm{Exp}_{y\in N(r)}\left[F(y)\right]-\mu\right|\geq\frac{\mu}{2}$$

כלומר - לכל פונקציה F על L, המכתיבה ממוצע μ על L, $(\beta,\lambda(\beta))-sampler$ מבטיח לנו שיש לכל היותר $\lambda(\mu)\cdot|R|$ צמתים $r\in R$ כך שהממוצע על F של שכני r (אלו איברים מ L) רחוק מ μ ביותר מ $\frac{\mu}{2}$.

נשים לב, כי כש F פונקציה בינארית, כלומר $F:L\to\{0,1\}$, הגרף מבטיח שלכל צומת ב- R מלבד חלק בגודל יחסי $\lambda(\mu)$ יש מספר שכנים מתוך $\{x:F(x)=1\}$ קרוב לממוצע μ.

הממוצע μ הוא החלק היחסי של L ש"סימנתי" בעזרת F. ה $sampler$ מבטיח שלרוב (מלבד...) הצמתים ב R, יחס השכנים המסומנים הוא גם כן μ. (ליתר דיוק, קרוב אליו ב $\frac{\mu}{2}$).

And from our paper

<u>הוכחת למה 2.2 (במאמר שלנו)</u>

ההוכחה להלן חלקית, ובחלקה גם לא נכונה. המשך ההוכחה ינתן ע"י קרין.

יהיו V_0,V_1 זוג מרחבים ליניאריים אורתוגונאליים כלשהם מעל שדה F_q. נקבע $V_0=F_q^m,V_1=F_q^m$.

יהא $X\subseteq V_0+V_1$ תת קבוצה בגודל יחסי μ, כלומר $\frac{|X|}{|V_0+V_1|}=\mu$. ויהא $W\subset V_1$ תת מרחב ליניארי רנדומאלי של V_1.

צ"ל: $\Pr\left[\left|\frac{|(V_0+W)\cap X|}{|V_0+W|}-\mu\right|\geq\frac{\mu}{2}\right]\leq\frac{4q^2}{|W|\mu}$.

נסמן ב $1_{x_1},\ldots,1_{x_{|X|}}$ את האינדיקטורים המציינים עבור האיבר ה- i של X (בהינתן סדר כלשהו של איברי X):

$$1_{x_i}=\begin{cases}\frac{1}{|V_0+W|} & x_i\in V_0+W\\ 0 & else\end{cases}$$

מתקיים $\mu_i=E\left[1_{x_i}\right]=\frac{\Pr\left[x_i\in V_0+W\right]}{|V_0+W|}\overset{\approx Uniform}{=}\frac{1}{|V_0+W|}\cdot\frac{|V_0+W|}{|V_0+V_1|}=\frac{1}{|V_0+V_1|}$.

נסמן $X_\Sigma = \sum_{i=1}^{|X|} 1_{x_i}$. X_Σ מציין את יחס מספר איברי $V_0 + W$ שנמצאים גם ב X מתוך כלל האיברים ב $V_0 + W$.

נסמן $\mu_\Sigma = \sum_{i=1}^{|X|} \mu_i$. מתקיים:

$$\mu_\Sigma = \sum_{i=1}^{|X|} \mu_i = \sum_{i=1}^{|X|} \frac{1}{|V_0 + V_1|} = |X| \frac{1}{|V_0 + V_1|} = \mu$$

ומאי שוויון Chebyshev נקבל:

$$\Pr\left[|X_\Sigma - \mu_\Sigma| \geq \frac{1}{2}\mu_\Sigma \right] \leq \frac{4}{\mu_\Sigma}$$

$$\Pr\left[\left| \frac{|(V_0 + W) \cap X|}{|V_0 + W|} - \mu \right| \geq \frac{1}{2}\mu \right] \leq \frac{4}{\mu} \leq \frac{4q^2}{|W|\mu}$$

Sampler Graphs – 2.3

גרף דו-צדדי בי-רגולרי $G(L,R)$ יקרא $(\alpha,\beta)-sampler$ אם בהינתן תת קבוצה של L בגודל יחסי $\mu \geq \alpha$ יש לכל היותר $\beta|R|$ צמתים עבורם ההסתברות שהשכנים מ L נמצאים בתת קבוצה רחוקה מהחלק היחסי μ ב $\frac{\mu}{2}$. כלומר תמיד יהיו (אם $\beta \neq 1$) צמתים עם מספר שכנים בין $\frac{1}{2}\mu$ ל $\frac{3}{2}\mu$.

בהינתן עולם U נציג שני inclusion graphs:

- *Independent* - L הם כל תתי הקבוצות בגודל s_l של U, ו R הם כל תתי הקבוצות בגודל s_r כש $s_r = t \cdot s_l$ עבור $t > 1$ שלם.
- *Subspaces* – עבור $U = F_q^m$, L הם כל תתי המרחבים הלינאריים ה d_l מימדיים. R הם כל תתי המרחבים הלינאריים ה d_r מימדיים, כש $d_r = c \cdot d_l$ עבור $c > 1$ שלם.

למה 2.3

שני הגרפים לעיל הם $(\alpha,b)-samplers$ כש:

- *Independent* - $\beta=e^{-\Omega(\alpha t)}$ בתנאי ש $\frac{(t\cdot s_l)^2}{|U|}\le e^{-\Omega(\alpha t)}$ ו- $\alpha t\ge\Omega\left(\ln\frac{1}{\alpha}\right)$.
- *Subspaces* - $\beta=O\left(\frac{1}{\sqrt{\alpha q^{(c-1)d_l}}}\right)$ בתנאי ש $\alpha^{\frac{3}{2}}q^{(c-1)\frac{d_l}{2}}>10$.

הוכחת היות ה Independent גרף sampler

נתון: $\beta=e^{-\Omega(\alpha t)}$, $\frac{(t\cdot s_l)^2}{|U|}\le e^{-\Omega(\alpha t)}$ ו- $\alpha t\ge\Omega\left(\ln\frac{1}{\alpha}\right)$

צ"ל: לכל תת קבוצה $F\subseteq L$ בגודל יחסי $\mu\ge\alpha$ יש לכל היותר $\beta|R|$ צמתים ב R עם

$$\left|\Pr_{l\in N(r)}[l\in F]-\mu\right|\ge\frac{\mu}{2}$$

יהא $M=s_r$ גודל תתי הקבוצות בצד ימין. תהי $S_1,...,S_t$ חלוקה קבועה של הקבוצה $[M]$ ל t קבוצות, בגודל s_l כל אחת.

תהא $G=S_M$ קבוצת הפרמוטציות על $[M]$.

נניח סידור כלשהו על העולם, ואז ניתן להסתכל על כל תת קבוצה $B\subseteq U$ בגודל M כ $M-tuple$ מסודר. וניתן לאיברים אינדקסים מ $[M]$. נסמן עבור $S\subseteq[M]$ ב $S(B)$ את איברי B שהאינדקסים שלהם נמצאים ב S.

בהינתן S_i קבוע כלשהו, ניתן לבחור ב $\pi\in G$ אקראי, והוא ימפה את S_i לתת קבוצה רנדומלית (בהתפלגות אחידה) אחרת בגודל s_l. אז בהינתן S_i קבוע, אם נבחר תת קבוצה $B\subseteq U$ בגודל M אקראית, ו $\pi\in G$ אקראית, אנו מקבלים ש $(\pi S_i)(B)$ הוא תת קבוצה רנדומלית (בהתפלגות אחידה) של U בגודל s_l.

יהיה $L'\subseteq L$ תת קבוצה בגודל יחסי $\lambda\ge\alpha$, אז מהעיל אנו מקבלים:

$$\lambda=\Pr_{i\in[t],B\in R,\pi\in G}\left[(\pi S_i)(B)\in L'\right]=Exp_{B\in R,\pi\in G}\left[\Pr_{i\in[t]}\left[(\pi S_i)(B)\in L'\right]\right]$$

(כי ההתפלגות היא אחידה)

אז עבור $\pi\in G$, $B\in R$ תתי הקבוצות $(\pi S_1)(B),...,(\pi S_t)(B)$ מהוות $t-tuple$ מפולג יוניפורמית של תתי קבוצות בגודל s_l של U זרות בזוגות. זה כמעט אותו דבר כמו לבחור יוניפורמית $t-tuple$ של איברים מ L.

נניח נבחר $S_1',...,S_t'$ יוניפורמית מ L. הסיכוי שלזוג S_i',S_j' כלשהו יהיה חיתוך לא ריק הוא לכל היותר t^2 פעמים הסיכוי שלשתי תתי קבוצות בגודל s_l רנדומליות של U יש חיתוך לא ריק.

הסיכוי שלשתי תתי קבוצות בגודל s_l רנדומליות של U - Q_i,Q_j - יש חיתוך לא ריק חסום ע"פ מרקוב ע"י $\frac{s_l^2}{|U|}$.

(אי שוויון מרקוב - $\Pr\left[|X|\ge a\right]\le \frac{E[X]}{a}$)

אצלנו:

$$\Pr\left[\left|Q_i\cap Q_j\right|\ge 1\right]\le E\left[\left|Q_i\cap Q_j\right|\right]=\overbrace{\frac{s_l}{|U|}}^{\text{one group is chosen. now this is the prob of a single member belonging to the chosen group}}\cdot\overbrace{s_l}^{\text{there are } s_l \text{ independent events for choosing members to the new group (like iid indicators)}}$$

על כן המרחק הסטטיסטי בין ההתפלגויות על $S_1^{'},\dots,S_t^{'}$ הוא לכל היותר $\frac{(t\cdot s_l)^2}{|U|}$ שע"פ הנתון קטן שווה מ $e^{-\Omega(\alpha t)}$.

אם כך קיבלנו

$$\Pr_{B\in R,\pi\in G}\left[\left|\Pr_{i\in[t]}\left[(\pi S_i)(B)\in L'\right]-\lambda\right|\ge\frac{\lambda}{3}\right]\le\Pr_{S_1^{'},\dots,S_t^{'}}\left[\left|\Pr_{i\in[t]}\left[(\pi S_i)(B)\in L'\right]-\lambda\right|\ge\frac{\lambda}{3}\right]+\frac{(t\cdot s_l)^2}{|U|}$$

(קשר סטטיסטי בבחירה לפי B ו- π לבין בחירת $S_1^{'},\dots,S_t^{'}$. הנתון הוא כמה רחוקה ההסתברות שתת קבוצה רנדומלית תהיה ב L' מהגודל היחסי של הקבוצה.)

לפי אי שוויון Chernoff נקבל

$$\Pr_{S_1^{'},\dots,S_t^{'}}\left[\left|\Pr_{i\in[t]}\left[S_i^{'}\in L'\right]-\lambda\right|\ge\frac{\lambda}{3}\right]\le e^{-\Omega(\lambda t)}\le e^{-\Omega(\alpha t)}$$

אז עבור $p=e^{-\Omega(\alpha t)}+(t\cdot s_l)^2/|U|\le e^{-\Omega(\alpha t)}$ נקבל כי $\Pr_{B\in R,\pi\in G}\left[\left|\Pr_{i\in[t]}\left[(\pi S_i)(B)\in L'\right]-\lambda\right|\ge\frac{\lambda}{3}\right]\le p$. לפי מיצוע נקבל כי לכל היותר ל $\sqrt{p}$ מהקבוצות $B\in R$, היו $\sqrt{p}$ פרמוטציות $\pi\in G$ כך שהחלק של $(\pi S_i)(B)$ שנופל לתוך L' היה גדול מ λ ביותר מ $\frac{\lambda}{3}$.

אזי המשלים הוא $1-\sqrt{p}$ קבוצות $B\in R$ שעבורן היו $1-\sqrt{p}$ פרמוטציות $\pi\in G$ כך שהחלק של $(\pi S_i)(B)$ שנופל לתוך L' הוא בין $\frac{2}{3}\lambda$ ל- $\frac{4}{3}\lambda$.

לסיום, עבור $B\in R$ נתון, ההסתברות שתת קבוצה רנדומלית בגודל s_l של B, נופלת לתוך L' היא $\mathrm{Exp}_{\pi\in G}\left[\Pr_{i\in[t]}\left[(\pi S_i)(B)\in L'\right]\right]$. מהעיל, בשביל כל למעט חלק $\sqrt{p}$ קבוצות B, הממוצע הזה מעל $\pi\in G$ יהיה לפחות $\left(1-\sqrt{p}\right)\frac{2}{3}\lambda\ge\frac{\lambda}{2}$ ולכל היותר $\frac{4}{3}\lambda+\sqrt{p}\le\frac{3}{2}\lambda$.

(בגלל ש $\sqrt{p}\le\frac{\lambda}{10}$ מתוך הנתון/הנחה $\alpha t\ge\Omega\left(\ln\frac{1}{\alpha}\right)$).

הסבר קצר: בשביל לחסום מלמטה, לקחנו את המקרה הכי קטן בו הקבוצה ה"רעה" של $\pi\in G$ תורמת 0 (הכי מעט שהיא יכולה), ובשביל לחסום מלמעלה, לקחנו את המקרה הכי גדול, בו הקבוצה הרעה של $\pi\in G$ תורמת 1 (הכי הרבה שהיא יכולה), ולכן מופיע $+\sqrt{p}$ בביטוי השני ולא בראשון.

מש"ל.

<u>הוכחת היות ה Subspaces גרף sampler</u>

הוכחה דומה לקודמת.

נגדיר $D = d_r$. עבור $t = \frac{q^D - 1}{q^{d_l} - 1}$ יהיו $S_1, ..., S_t$ אוסף קבוע של תתי מרחב לינאריים מסדר d_l של F_q^D, הזרים בזוגות למעט וקטור ה- 0. האוסף מובטח לנו מלמה 2.1.

G זה אוסף כל המטריצות $D \times D$ הלא סינגולריות.

עבור תת מרחב לינארי מסדר d_l S, ותת מרחב לינארי מסדר D B, נסמן ב- $S(B)$ את התת מרחב המתאים ב B.

כמו בהוכחה הקודמת, דנים שעבור $B \in R$ ו $A \in G$ רנדומליים, המרחב $(AS_i)(B)$ מתפלג יוניפורמית מעל L.

הסיכוי לבחור 2 תתי מרחב מסדר d_l <u>תלויים לינארית</u> של U הוא $\frac{q^{2d_l}}{q^m}$. למה? כי נניח בחרתי אחד אקראי מ L. מה הסיכוי שאבחר בתת מרחב תלוי לינארית? זה כמות המרחבים שתלויים לינארית בזה שבחרתי חלקי כמות כל המרחבים. יש q^{d_l} מרחבים תלויים לינארית בזה הנבחר, שכן תת מרחב מסדר d_l מוגדר ע"י d_l וקטורים בת"ל. כל מרחב המכיל כפולה בסקלר של אחד הוקטורים יהיה תלוי לינארית בשלנו, אזי צריך לבחור איזה סקלר מתוך q d_l פעמים – q^{d_l}. נחלק בכמות המרחבים הלינאריים $\frac{q^m}{q^{d_l}}$ ונקבל $\frac{q^{2d_l}}{q^m} = \frac{q^{d_l}}{\frac{q^m}{q^{d_l}}}$. (כשלכל אורך הדיון הנחתי כי $m = D$ - לא מצאתי התייחסות במסמך).

מכיוון שהסיכוי שקיבלנו זניח, נגיד ש $\left((AS_i)(B), (AS_j)(B)\right)$ מתפלג אחיד מעל L ולא רק מעל המרחבים הבת"ל.

אזי כמו קודם מתקיים:

$$\lambda = \Pr_{B \in R, A \in G, i \in [t]}\left[(AS_i)(B) \in L'\right] = \mathrm{Exp}_{A \in G, B \in R}\left[\Pr_{i \in [t]}\left[(AS_i)(B) \in L'\right]\right]$$

ולפי אי"ש צ'בישב:

$$\Pr_{B \in R, A \in G}\left[\left|\Pr_{i \in [t]}\left[(AS_i)(B) \in L'\right] - \lambda\right| \geq \frac{\lambda}{3}\right] \leq \frac{\mathrm{Var}_{B \in R, A \in G}\left[\sum_{i=1}^{t} \chi\left[(AS_i)(B) \in L'\right]\right]}{\frac{(t\lambda)^2}{9}}$$

נסביר: אי"ש צ'בישב נתון ע"י :

$$\Pr\left[\left|X - E[X]\right| \geq \mu\sqrt{V[X]}\right] \leq \mu^{-2}$$

אצלנו:

$$\Pr_{B\in R,A\in G}\left[\left|\Pr_{i\in[t]}\left[\left(AS_i\right)\left(B\right)\in L'\right]-\lambda\right|\geq\frac{\lambda}{3}\right]$$
$$=\Pr_{B\in R,A\in G}\left[\left|\Pr_{i\in[t]}\left[\left(AS_i\right)\left(B\right)\in L'\right]t-\lambda t\right|\geq\frac{\lambda t}{3}\right]$$
$$=\Pr_{B\in R,A\in G}\left[\left|X-E\left[X\right]\right|\geq\frac{\lambda t}{3}\right]\Rightarrow\mu=\frac{\lambda t}{3\sqrt{V\left[X\right]}}\Rightarrow$$
$$\Pr_{B\in R,A\in G}\left[\left|\Pr_{i\in[t]}\left[\left(AS_i\right)\left(B\right)\in L'\right]-\lambda\right|\geq\frac{\lambda}{3}\right]\leq\mu^{-2}=\frac{9V\left[X\right]}{\left(\lambda t\right)^2}$$

מכאן ניתוח השונות ברורה במסמך, ומובילה אותנו לתוצאה:

$$\Pr_{B\in R,A\in G}\left[\left|\Pr_{i\in[t]}\left[\left(AS_i\right)\left(B\right)\in L'\right]-\lambda\right|\geq\frac{\lambda}{3}\right]\leq\frac{18}{\lambda t}$$

ועתה כמו בסיום ההוכחה הקודמת, לפי שיקולי מיצוע מגדירים $p=\frac{18}{\lambda t}$, ואומרים שלכל ה- B -ים למעט חלק בגודל יחסי $\sqrt{p}$ מתקיים שלכל $A\in G$ למעט חלק בגודל יחסי $\sqrt{p}$ החלק של תתי המרחב כך ש $\left(AS_i\right)\left(B\right)\in L'$ הוא בין $\frac{2}{3}\lambda$ ו $\frac{4}{3}\lambda$.

לסיום, עבור $B\in R$ נתון, ההסתברות שתת קבוצה רנדומלית בגודל s_l של B , נופלת לתוך L' היא $\mathrm{Exp}_{\pi\in G}\left[\Pr_{i\in[t]}\left[\left(\pi S_i\right)\left(B\right)\in L'\right]\right]$. מהעיל, בשביל כל למעט חלק $\sqrt{p}$ קבוצות B , הממוצע הזה מעל $A\in G$ יהיה לפחות $\left(1-\sqrt{p}\right)\frac{2}{3}\lambda\geq\frac{\lambda}{2}$ ולכל היותר $\left(1-\sqrt{p}\right)\frac{4}{3}\lambda\leq\frac{4}{3}\lambda+\sqrt{p}\leq\frac{3}{2}\lambda$.

(בגלל ש $\sqrt{p}\leq\frac{\lambda}{10}$ מתוך הנתון/הנחה $\lambda^{\frac{3}{2}}q^{(c-1)\frac{d_l}{2}}>10$).

מש"ל.

2.4 תכונות של Samplers

למה 2.4 – עבור $G\left(L,R\right)$ שהוא $\left(\alpha,\beta\right)-sampler$, נקבע $\lambda\geq0,\rho\leq1$ כך ש $\lambda\geq\alpha$ ו- $\frac{\lambda\rho}{10}\geq\beta$. אזי לכל תת קבוצה $L'\subseteq L$ בגודל יחסי λ ולכל $R'\subseteq R$ בגודל יחסי ρ מתקיים:

$$\left|\Pr_{r\in R',l\in N(r)}\left[l\in L'\right]-\lambda\right|\leq\frac{2}{3}\lambda$$

צד שמאל של האי"ש הוא לכל היותר $Exp_{r\in R'}\left[\left|\Pr_{l\in N(r)}\left[l\in L'\right]-\lambda\right|\right]$ (למה?). מתוך ההגדרה של $sampler$ אנחנו מקבלים שעבור כל פרט לחלק $\frac{\beta}{\rho}$ של צמתים $r\in R'$ מתקיים $\left|\Pr_{l\in N(r)}\left[l\in L'\right]-\lambda\right|\leq\frac{1}{2}\lambda$. אז התוחלת הכוללת על $r\in R'$ יהא לכל היותר $\frac{\lambda}{2}+\frac{\beta}{\rho}\leq\frac{\lambda}{2}+\frac{\lambda}{10}$.

למה 2.5 – ההגדרה של *sampler* אינה סימטרית. היא רוצה דגימה טובה של צמתים מ- L ע"י צמתים מ- R. למה 2.5 אומרת שעבור קבוצות גדולות מ- R אנו מקבלים תכונה דומה סימטרית. פרטים במסמך.

ההוכחה מחלקת את L ל bad_1 ו- bad_2 שאלו החלקים בהם היא מקבלת יותר או פחות בהתאמה מהממוצע. מראים שהגודל שלהם בהכרח קטן מ λ כל אחד ע"י סתירה של הגדרתם מול הגדרת ה sample. על כן סה"כ יש חלק יחסי 2λ צמתים "רעות" מ L.

(ההמשך כבר פחות מפורט, ומסתמך בעיקר על המסמך. לכן חלק מהסימנים חסרי הגדרה)

למה 2.6 – מחדדת את למה 2.5 עבור גרף ה *independent*.

מסקנה 2.7 – השלכת למה 2.6 על המקרה הכללי של *inclusion graph*. מקבלים כי יש לכל היותר חלק יחסי של $O\left(\left(\log\frac{1}{\rho}\right)\Big/\left(\frac{k}{k'}\right)\right)$ של צמתים מ L עבורם $\left|\Pr_{r\in N(l)}\left[r\in R'\right]-\rho\right|\le \nu\rho$.

למה 2.8 – ע"י חסם Chernoff-Hoeffding, בהנתן $S\subseteq U$ בגודל יחסי λ ו $0<\nu<1$ קבוע. אזי גדלי תתי הקבוצות של U^k עבורם החיתוך עם S גדול ב ν או קטן בו מהממוצע חסום ע"י $e^{-\Omega(\lambda k)}$.