Hardness of Maximum Independent Set in Structured Hypergraphs

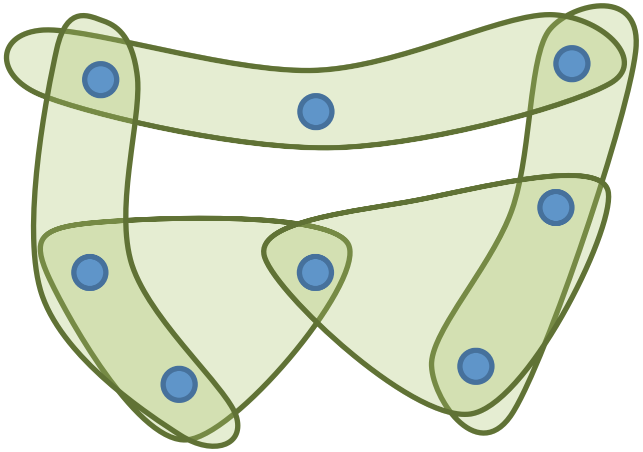
Rishi Saket

IBM TJ Watson

Joint works with: Subhash Khot (NYU)

Sushant Sachdeva (Berkeley)

Independent Set in Hypergraphs



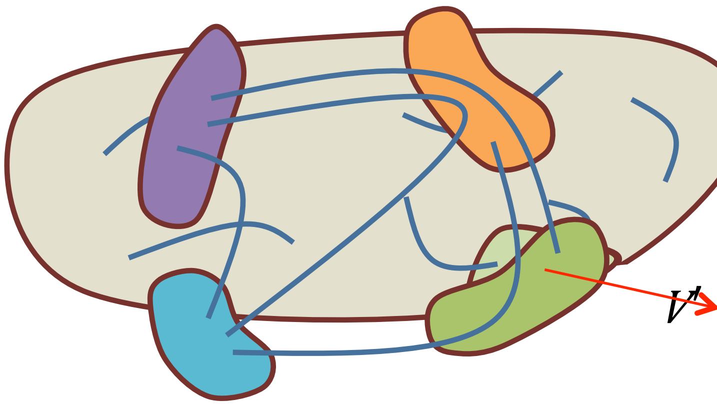
Independent Set : Subset of vertices not containing all the vertices of any hyperedge.

k-colorable : Vertices can be partitioned into k disjoint independent sets.

Maximum Independent Set

In hypergraphs with specific *structural* property

Almost k-Colorability



G has Indep. Set of $\approx n/k$ size.

$G(V,E)$

G' : remove V' and all $e \sim V'$

k-colorable

Hypergraph $G(V,E)$ on n vertices is almost k-colorable if, $\exists\, V' \subseteq V, |V'| \leq \varepsilon n$ s.t. G' is k-colorable.

Maximum Independent Set in (almost) 2-colorable hypergraphs

NP-hard to find independent set of relative size:

- $(\log\log n / \log\log\log n)^{-1}$ in 2-colorable 4-uniform hypergraphs [GHS 00] [Hol 02]
- (Assuming UGC) any constant $\delta > 0$ in,
  - Almost 2-colorable 3-uniform hypergraphs, [Guruswami Sinop 10]

Our Results [Khot S. 13]

NP-hard to find independent set of relative size:

$2^{-(\log n)^{1-\varepsilon}}$ in almost 2-colorable 4-uniform hypergraphs.

Any constant $\delta > 0$ in almost 2-colorable 3-uniform hypergraphs, *without assuming UGC*.

Any constant $\delta > 0$ in 2-colorable 3-uniform hypergraphs, assuming d-to-1 Games Conjecture.

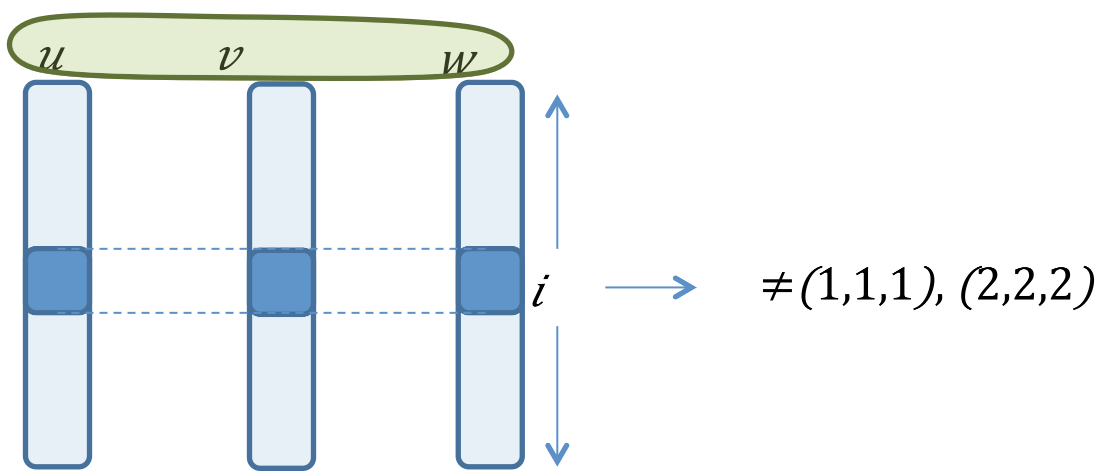
A "Dictatorship" Gadget

- Toy instance: 3-uniform hypergraph over Label Set L.
  - Yes Case: Every $i \in L \Rightarrow$ "good" solution, i.e. an almost 2-coloring.
  - No Case: Every large independent set ⇒ "small" subset of L.

Ground set $\Omega = \{*, 1, 2\}$. Measure $\mu(1) = \mu(2) = 1 - \varepsilon/2$, $\mu(*) = \varepsilon$.

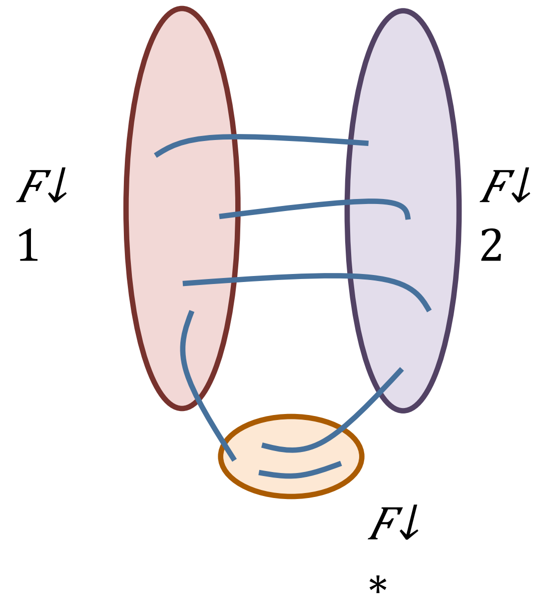
Vertices $V := \Omega^L$ with weights $wt(v) := \mu^L(v)$ for $v \in \Omega^L$.

Hyperedges E : $\{u, v, w\} \in E \Leftrightarrow \forall i \ (u_i, v_i, w_i) \neq (1,1,1), (2,2,2)$



YES Case

- Fix any $i \in L$.
- Define:
  - $F_1 = \{u \mid u_i = 1\}$,
  - $F_2 = \{u \mid u_i = 2\}$,
  - $F_* = \{u \mid u_i = *\}$



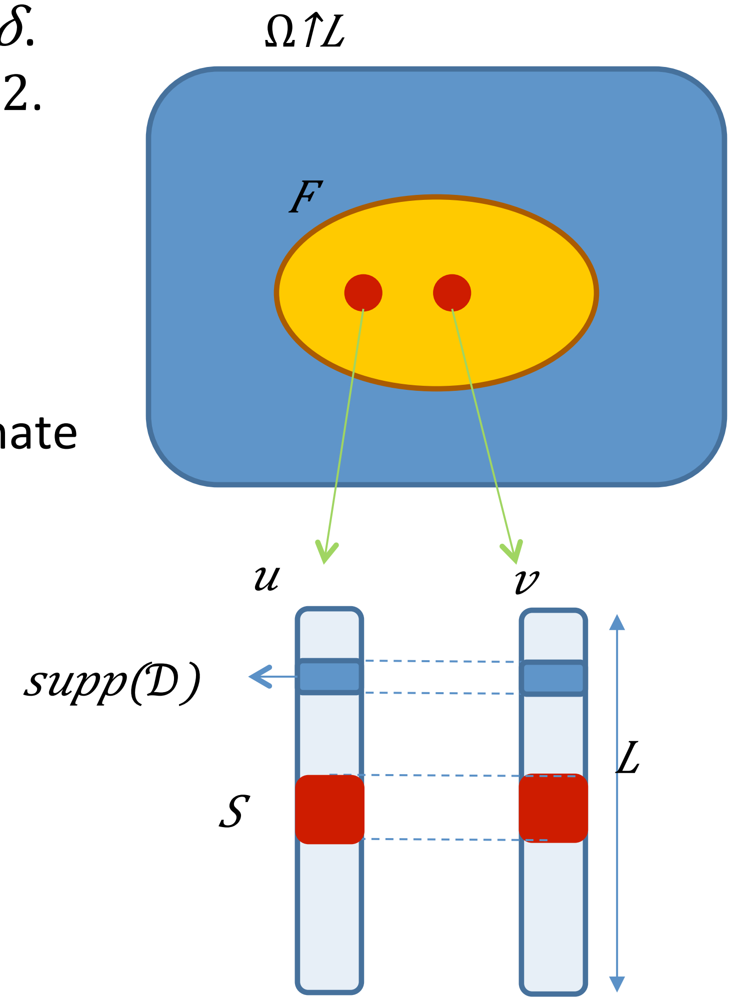
- F_1 and F_2 are disjoint independent sets of weight $1-\varepsilon/2$ each.
  - $u,v,w \in F_1 \Rightarrow (u_i, v_i, w_i) = (1,1,1)$
  - $u,v,w \in F_2 \Rightarrow (u_i, v_i, w_i) = (2,2,2)$
- "Good" solution corresponding to each $i \in L$.
  - Removing F_* + incident hyperedges makes it 2-colorable.

No Case

- Let $F \subseteq \Omega^{L}$ be maximal Ind. Set. $wt(F) \geq \delta$.
  - Monotone : Closed under changing $*$ to 1,2.
- Goal : "Decode" F to small subset of L.

- Let $\mathcal{D}$ be a distn. on Ω^{2} :
  - Sample uniformly from $\{(1,2),(2,1)\}$
  - Independently w.p. ε change each coordinate to $*$.
  - $(1,1), (2,2) \notin supp(\mathcal{D})$
  - Both marginals of $\mathcal{D}$ are identical to μ.

- <u>Key Property</u>: There exist $u, v \in F$ and $S \subseteq L$ such that:
  - $\forall i \in L - S,\ (u_{i}, v_{i}) \in supp(\mathcal{D})$
  - $|S| \leq \exp(poly(1/\delta, 1/\varepsilon))$

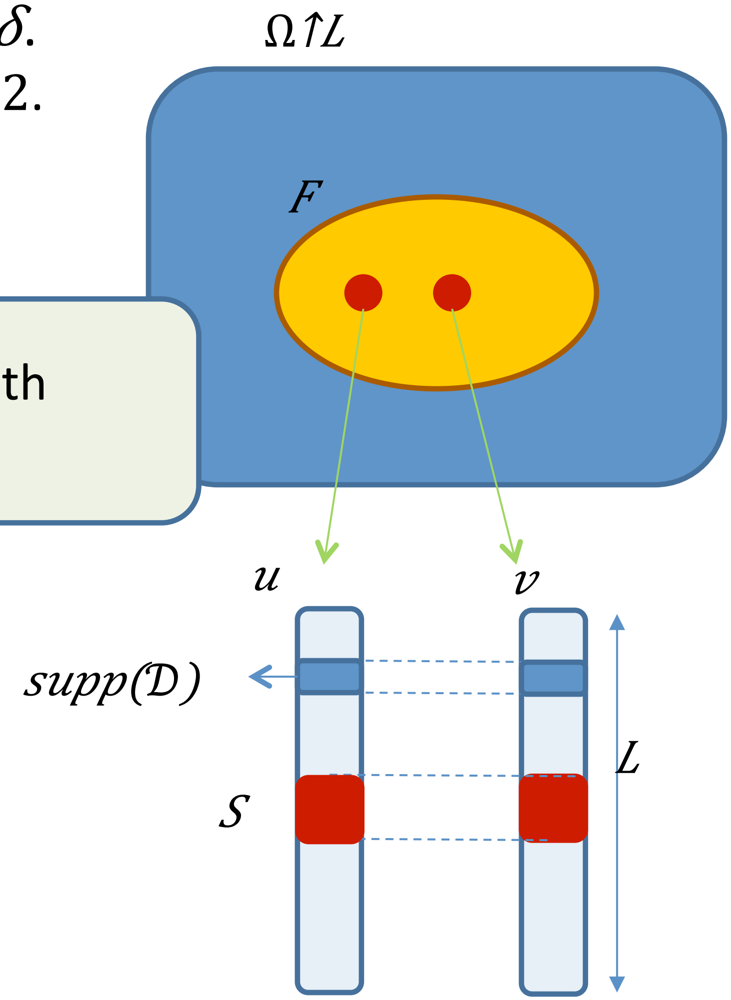


No Case

- Let $F \subseteq \Omega^L$ be maximal Ind. Set. $wt(F) \geq \delta$.
  - Monotone : Closed under changing $*$ to 1,2.
- Goal : "Decode" F to small subset of L.

- Let $\mathcal{D}$ be a distn. on Ω^2 :
  - Sampl
  - Indepe
    to $*$.
  - $(1,1), (2,2) \notin supp(\mathcal{D})$
  - Both marginals of $\mathcal{D}$ are identical to μ.

Combine Dictatorship Gadget with Multilayered PCP [DGKR 03]



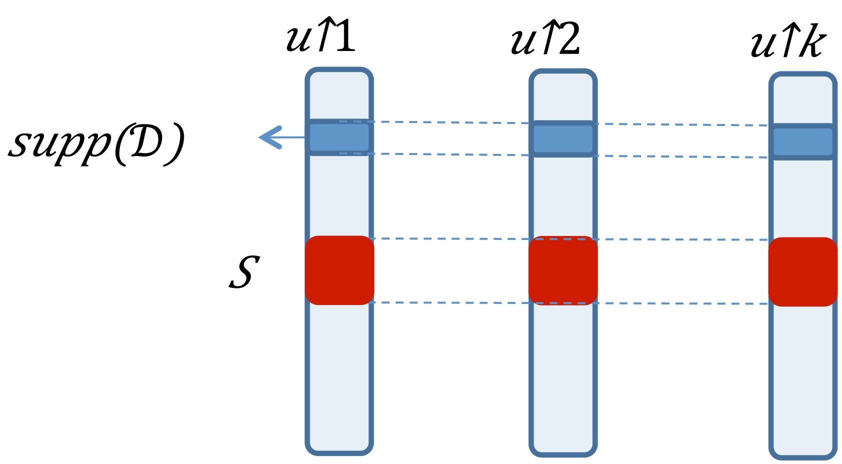
- Key Property: There exist $u, v \in F$ and $S \subseteq L$ such that:
  - $\forall i \in L - S,\ (u_i, v_i) \in supp(\mathcal{D})$
  - $|S| \leq \exp(poly(1/\delta, 1/\varepsilon))$

Proof Techniques

- $F \subseteq \Omega^{L}$ be monotone with $wt(F) \geq \delta$,
- [Russo] F has *avg. sensitivity* $O(1/\varepsilon)$.

- [Friedgut, Sachdeva Tulsiani 11] F is very close to a *junta* F' on $S \subseteq L$.
  - $|S| \leq \exp(poly(\varepsilon,\delta))$.

- [Dinur Safra 02, DKPS 10] Distn $\mathcal{D}$ on Ω^{k} with marginals identical to μ.
  - $\exists\, u^{1},\ldots,u^{k} \in F$ s.t. $(u_i^{1},\ldots,u_i^{k}) \in supp(\mathcal{D})$ for $i \in L - S$.
  - Jointly sample "extensions" of junta F' using $\mathcal{D}$.

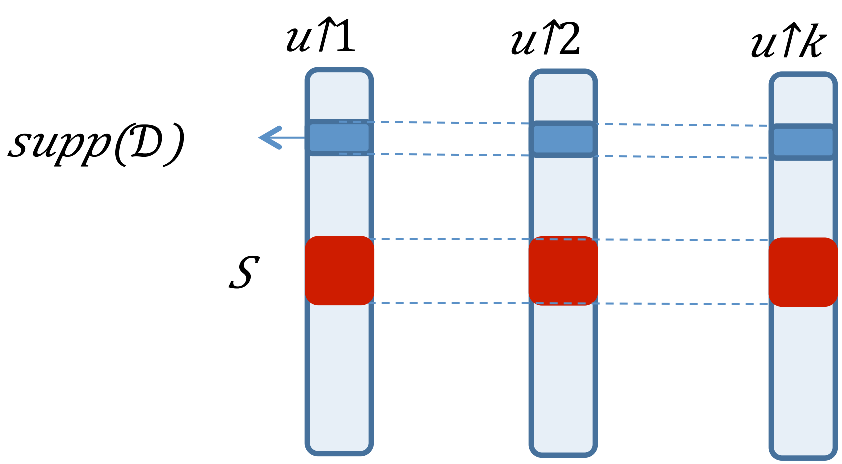


Proof Techniques

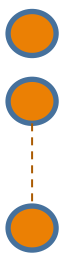
- $F \subseteq \Omega^L$ be monotone with $wt(F) \geq \delta$, $\mu(*) = \varepsilon$.
- [Russo] F has *avg. sensitivity* $O(1/\varepsilon)$.

- [Friedgut, Sachdeva Tulsiani 11] F is very close to a *junta* F' on $S \subseteq L$.
  – $|S| \leq \exp($
- [Dinur Safra als identical to μ.
  – $\exists\ u_1, \ldots$ $L - S$.
  – Jointly sample "extensions" of junta F' using $\mathcal{D}$.

Other problems where UGC can be avoided using similar techniques ?



Hardness of Scheduling Problems and Hypergraph Independent Set [Bansal Khot 10]



Job i:
Weight w_i
p_{ij} time on machine j

N jobs

M machines

<u>Concurrent Open Shop</u>
Minimize sum of weighted completion times. 2-approx. [GKP 07]

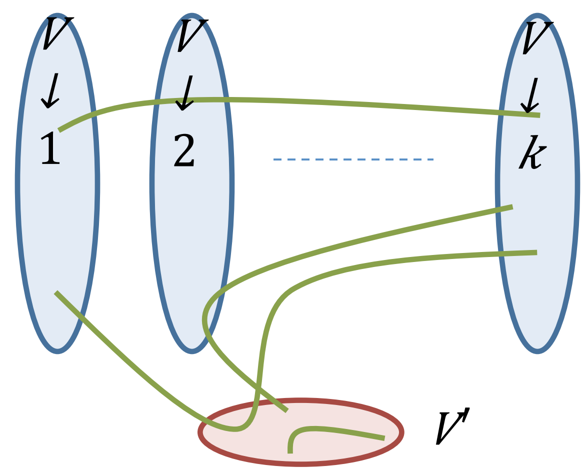
$(2-\varepsilon)$-UGC hard to approx. [BK 10]

<u>Ind. Set. in almost k-partite k-uniform hypergraphs</u> [BK 10]

Given a k-uniform hypergraph $G(V, E)$, UGC-hard to decide whether:

- Yes Case: V can be partitioned into $V', V_1, \ldots, V_k$ s.t. $|V'| = \varepsilon|V|$, $|V_1| = \ldots = |V_k|$ and each $e \in E$ has at most one vertex from any V_i.

- No Case: No Ind. Set of size $\varepsilon|V|$.



(Almost k-partite)

Hardness of Scheduling Problems and Hypergraph Independent Set [Bansal Khot 10]

Job i:
Weight w_i
p_{ij} time on machine j

N jobs

Concurrent Open Shop

Minimize sum of weighted completion times. 2-approx. [GKP 07]

$(2-\delta)$ UGC hard to approx. [BK 10]

[Sachdeva S. 13] Optimal $(2-\delta)$ NP-hardness for Concurrent Open Shop *without assuming UGC.*

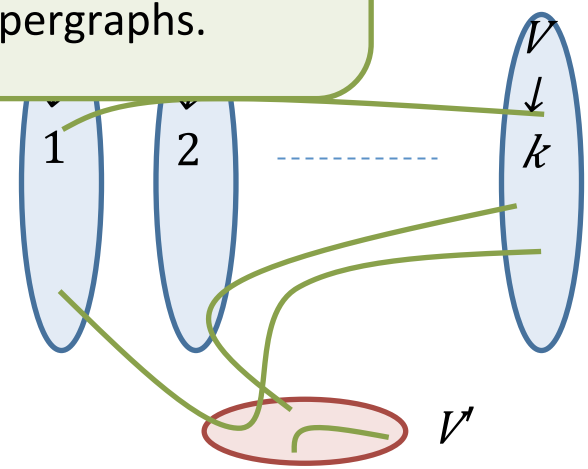
Indep. Set in "k-blocked-partite" hypergraphs.

Ind. Set in … hypergraph

Given a k-uniform hypergraph $G(V, E)$, UGC-hard to decide whether:

- Yes Case: V can be partitioned into V', V_1, …, V_k s.t. $|V'| = \varepsilon|V|$, $|V_1| = \ldots = |V_k|$ and each $e \in E$ has at most one vertex from any V_i.

- No Case: No Ind. Set of size $\varepsilon|V|$.



(Almost k-partite)

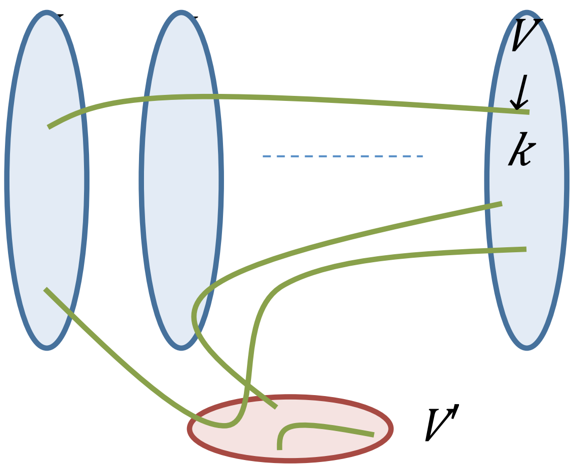
Bypassing UGC [Sachdeva S. 13]

Independent Set in "k-blocked partite" hypergraphs.

Given k^2 -uniform hypergraph $G(V,E)$, it is NP-hard to decide:

Yes Case

- V can be partitioned into V' ,V_1 ,…,V_k s.t. $|V'|=\varepsilon|V|$, $|V_1|=\ldots=|V_k|$.
- Any $e\in E$ can be divided into k blocks of k vertices each s.t. at least $k-1$ blocks are "good" : contain at most one vertex from any V_i .



No Case : No Indep. Set. of size $\varepsilon|V|$.

Reduction to Concurrent Open Shop

- Machines : Hyperedges E
- Jobs: Vertices V , unit weight.
  - $v \in V$ has unit processing time on each $e \sim v$.

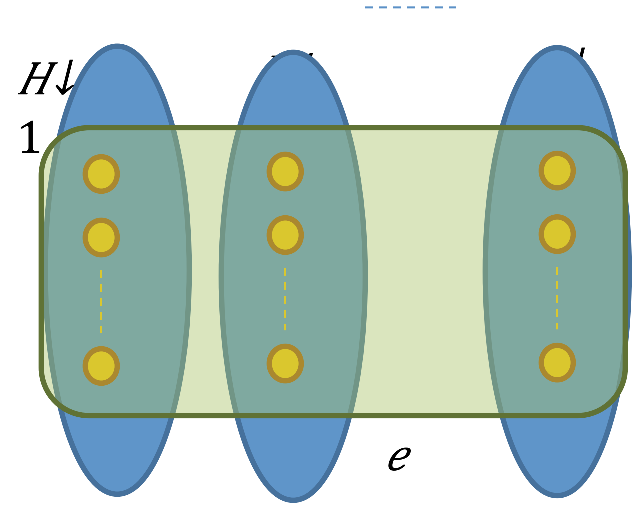
- Yes Case : Each e processes its k^2 jobs according to $V_1, V_2, \ldots, V_k, V'$.
  - Jobs V_i completed by time $ik+k$.
  - Jobs V' completed by time k^2 .
  - Average completion time : $\approx k(k+1)/2 + \varepsilon k^2$.

- No Case : Set of jobs completed by time $k^2 - 1$ is an indep. set.
  - Average completion time : $\geq (1-\varepsilon) k^2$.

- Choose large k and small ε : $(2-\delta)$- NP-hardness factor.

"Dictatorship" Gadget

Instance : k^2-uniform hypergraph over L.

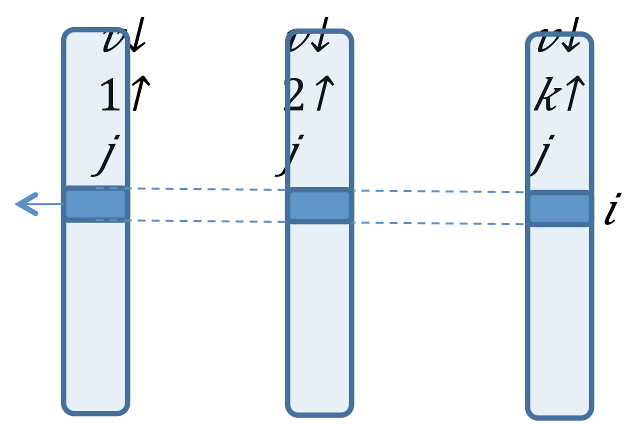


Ground set $\Omega=\{*,1,2,\dots,k\}$.
$\mu(*)=\varepsilon, \mu(1)=\mu(2)=\dots=\mu(k)=1-\varepsilon/k$.

Vertices $V:=\bigcup_{j=1}^{k} H_j$ where H_j is copy of Ω^L .

Hyperedges E: Select k vertices from each H_j such that:
for each $i\in L$, for $k-1$ of the H_j's :

Contains at most one value of r for each $r\in[k]$



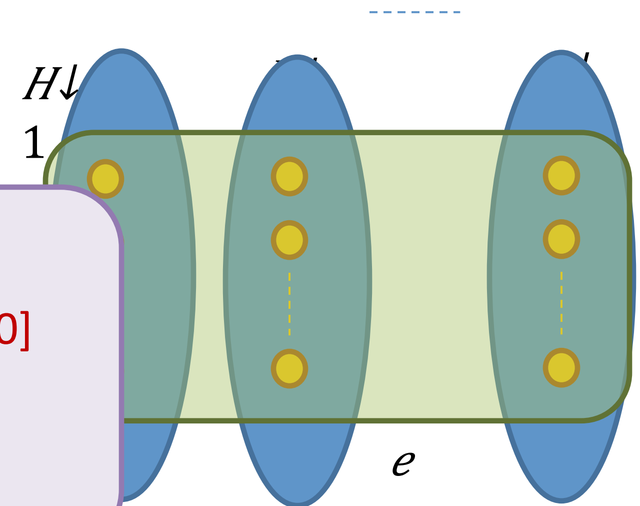
“Dictatorship” Gadget

Instance : $k\uparrow 2$ -uniform hypergraph over L.

Ground set $\Omega=\{*,1, 2, \ldots,k\}$.
$\mu(*)=\varepsilon, \mu(1$

Vertices
$\Omega\uparrow L$.

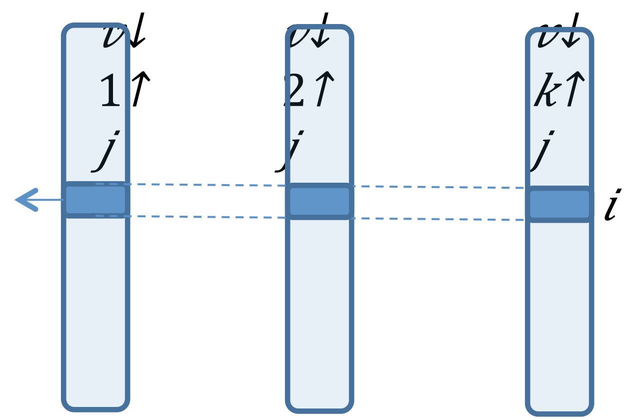
Hyperedg
such that:
for each $i\in L$, for $k-1$ of the $H\downarrow j$’s :



Analysis follows using:
[Russo] [Friedgut, ST 11] [DS 02, DKPS 10]

Combine with variant of Label Cover

Contains at most one value of r for each $r\in[k]$



Conclusions

- NP-hardness (bypassing UGC) results for Max. independent set in
  - Almost 2-colorable 3-uniform hypergraphs,
  - "k-blocked partite" hypergraphs.
  - Optimal $(2-\delta)$-inapprox. for Concurrent Open Shop, Assembly Line Scheduling.

- Future Directions.
  - Prove indep. set results for exactly colorable hypergraphs ?
  - Invariance Principle with Label Cover for better results for Vertex Cover, Independent Set …?
  - Scheduling with precedence constraints. $(2-\delta)$-UGC hard [BK 09]
  - Use LP integrality gaps for NP-hardness ? [KMTV 11]

Thank You!