Optimal Hardness Results for "somewhat" Hard Problems

Young Kun Ko

Princeton

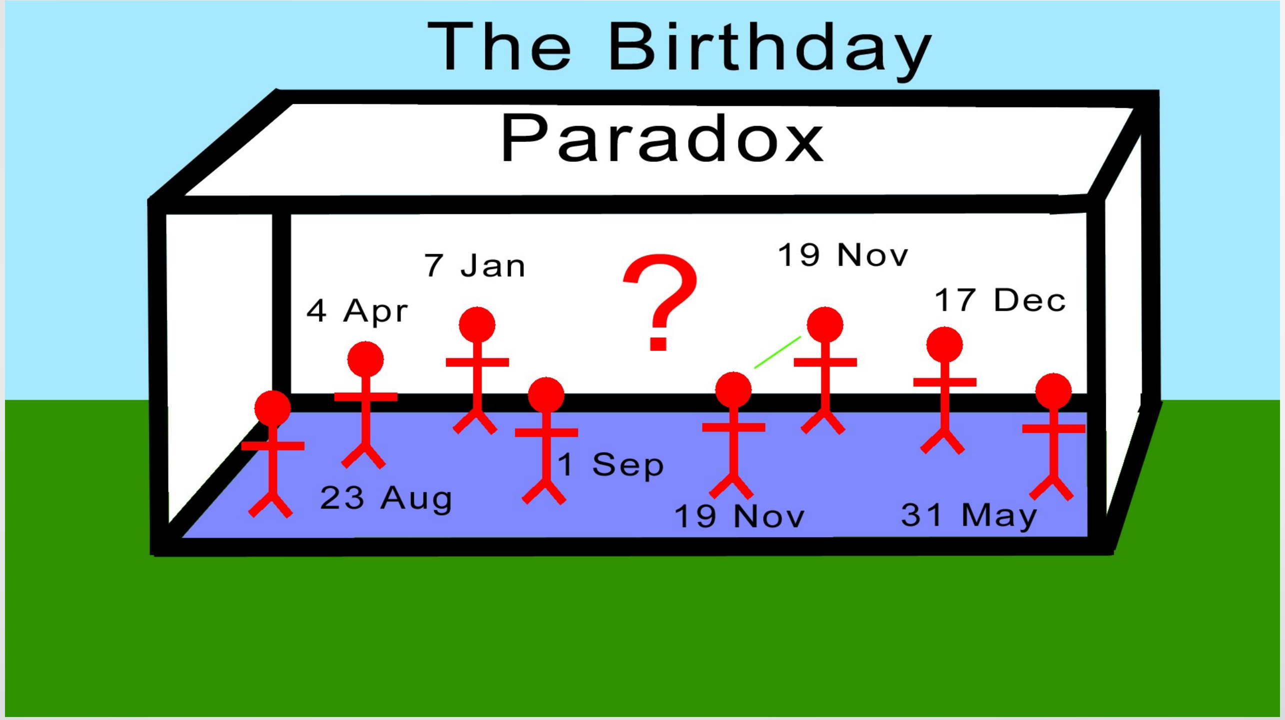
China Theory Week 2016

Table of Contents

Brief Intro to Birthday Repetition

Based on [Aaronson Impagliazzo Moshkovitz 2014]

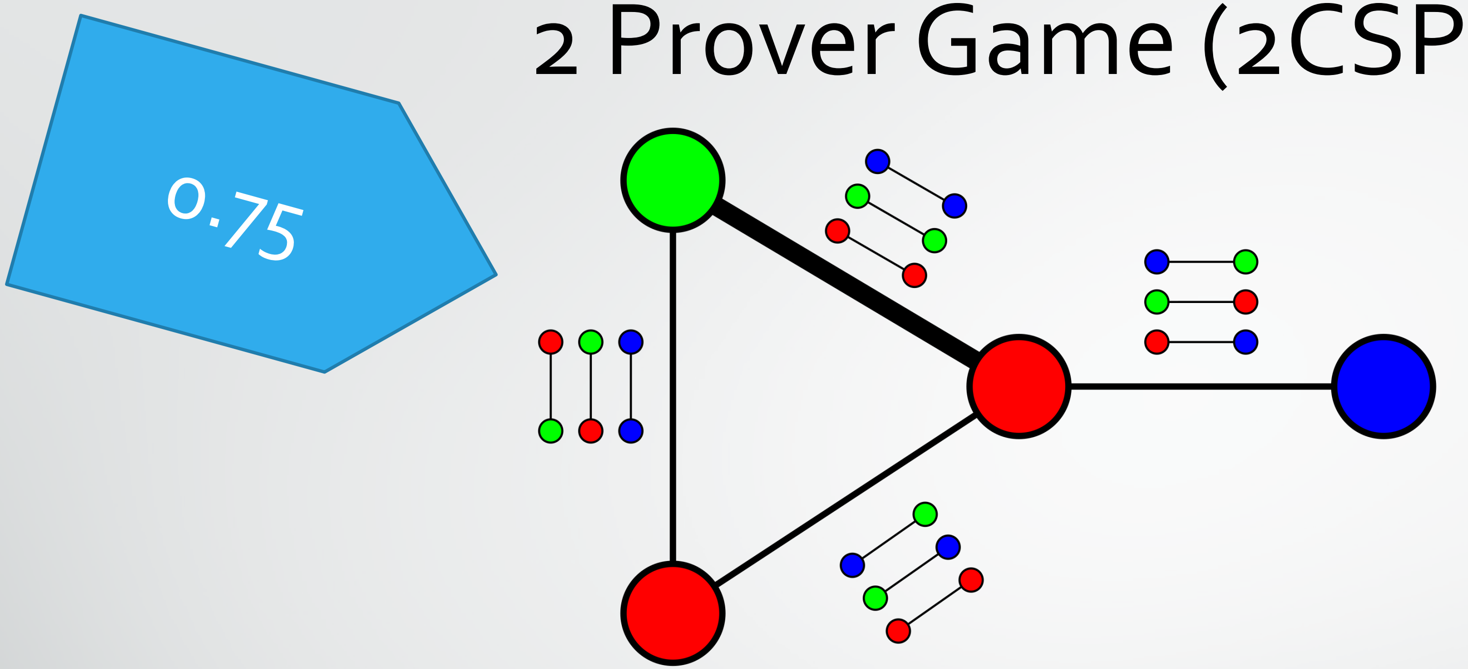
Birthday Paradox



- Given random two sets of size $\tilde{\Omega}(\sqrt{n})$, from the universe of size n there exists an intersection w.h.p !!

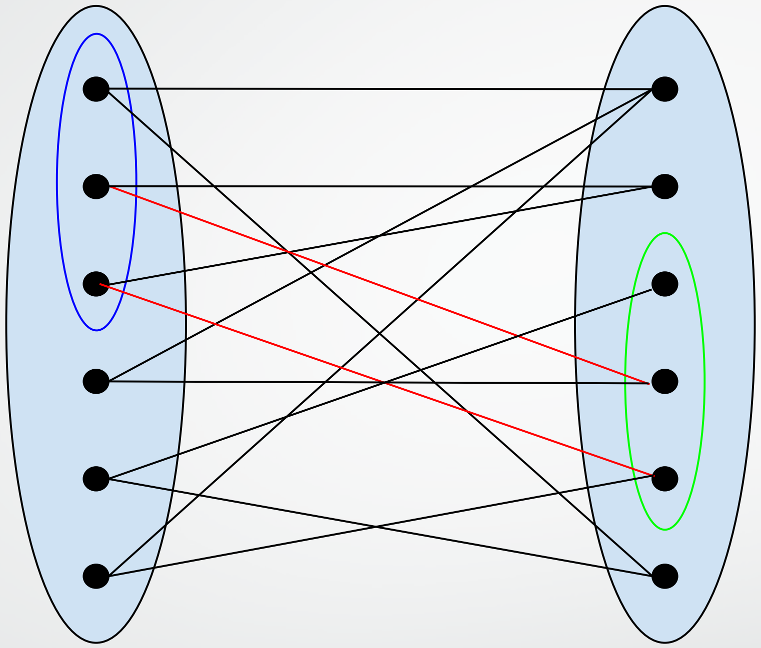
Proof) $\dfrac{\binom{n-\sqrt{n}}{\sqrt{n}}}{\binom{n}{\sqrt{n}}} \approx \left(1 - \dfrac{1}{\sqrt{n}}\right)^{\sqrt{n}} = \Theta(1)$

2 Prover Game (2CSP)



- Consider a game on a bipartite graph G , the referee picks an edge $(u, v) \in E$, give each side to Alice & Bob and asks for "coloring."
  Accepts if the coloring is valid for the picked edge.
- Given an instance, what is the best coloring and its "value"?
- Main Bottleneck on the hardness side : Alice and Bob's inputs are **correlated**
  - Provers can try to guess what the other side got as an input …

Getting Rid of the Correlation !



- **Birthday Repetition** Pick two random sets of challenge of size $\widetilde{\Omega}(\sqrt{n})$
- Then we can "approximately" check a random challenge
- Correlation no longer an issue due to the construction !
- Such games are called **"Free Games" :** $\boldsymbol{D_x \perp D_y}$

Main Intuition of [AIM'14]

Sub-exponential Size Reduction!

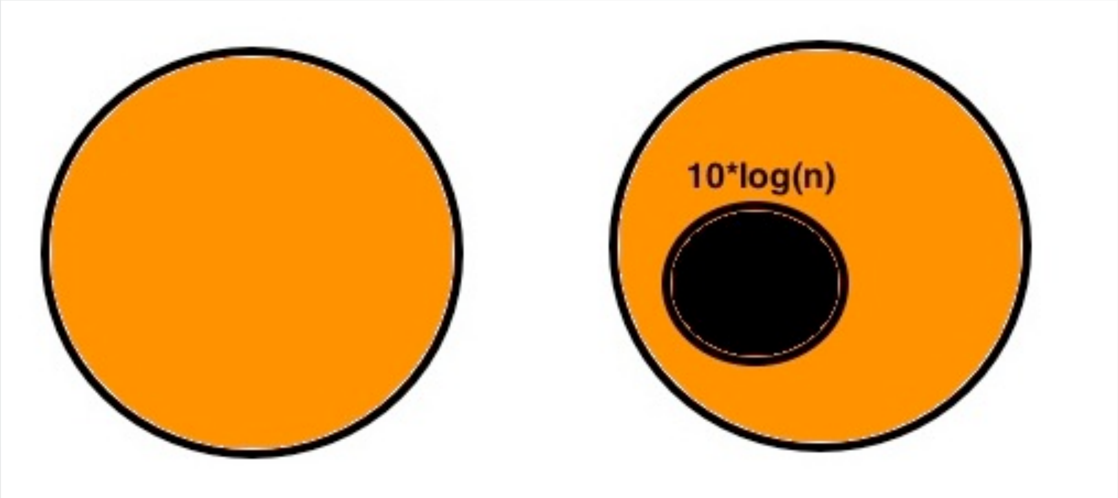
Proof ☑
Joke ☑

- **If $\boldsymbol{N = 2^{\sqrt{n}}, N^{\log N} = 2^n}$!**
- Sub-exponential Size reduction → Quasi-Polynomial Time Hardness assuming Exponential Time Hypothesis !
- Challenge as sets of size $\tilde{\Omega}(\sqrt{n})$ and assignments are local assignments on the sets
- Can start reduction from known NP-hard near-linear 2CSPs and preserve their value

Conjecture of Interest

- **Exponential Time Hypothesis** (Impagliazzo & Paturi '01)
  - Naïve Brute Force is the best you can do (**asymptotically**)
  - Difference with SETH ? (1.99^n not ruled out)
- **Planted Clique Conjecture** – Folklore ? (Target for replacement)

Planted Clique Conjecture



- Distinguish between the following two cases :
  - $G\left(n, \frac{1}{2}\right)$ vs. $G\left(n, \frac{1}{2}\right)$ planted with a clique of size $10 \log n$
- Search Version : find such planted clique ?
- Is Average case crucial ?? Lower-Bound ??
- Quasi-polytime algorithm obvious : the conjecture is that this is the best algorithm

Spectral Based poly time algorithm by Alon & Sudakov only works for $\sqrt{n}$-size clique

Summary of [AIM’14]

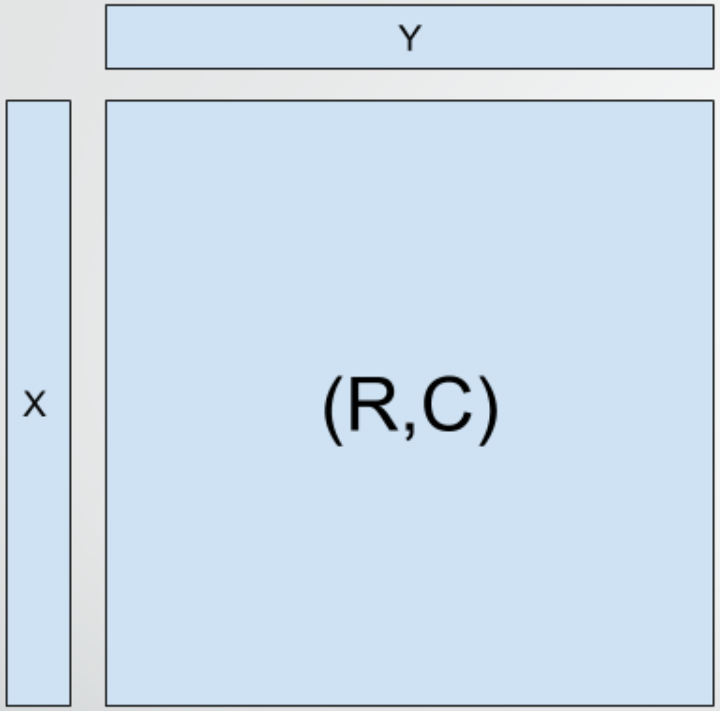
- Using Birthday Repetition, Can transform any given 2CSP (with some mild conditions) to free games
- Then starting from known NP-hard two player games, can show that ε-approximation to Free-game Value quasi-poly hard (under ETH) !
- Matching Upper Bounds known [AIM’14] [BH’13]

Application: Hitting Set Argument in 2-Player Game

Based on [Braverman **K** Weinstein' 15], [Cheng **K** '16]

Table of Contents

Approximate Nash Equilibrium



$$x^T R\, y > e_i^T R y + \epsilon$$

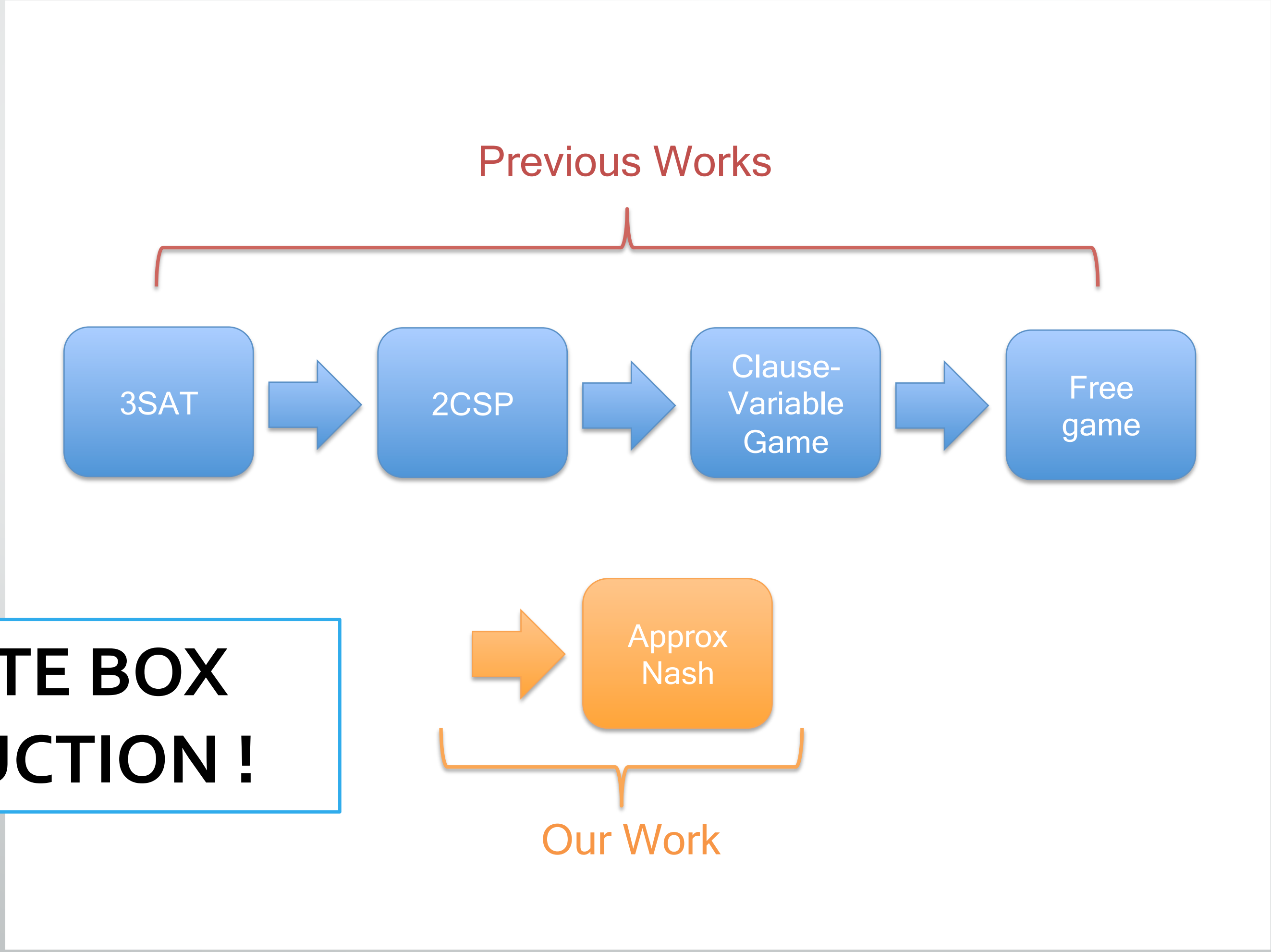
$$x^T C\, y > x^T C e_i + \epsilon$$

Good Nash) Total Payoff must be good !

- Alice, Bob plays a (strategic) game, represented by payoff matrix (R, C)
- Equilibrium → No player has incentive to deviate
- ϵ-approximate Nash → No player has more than ϵ incentive to deviate

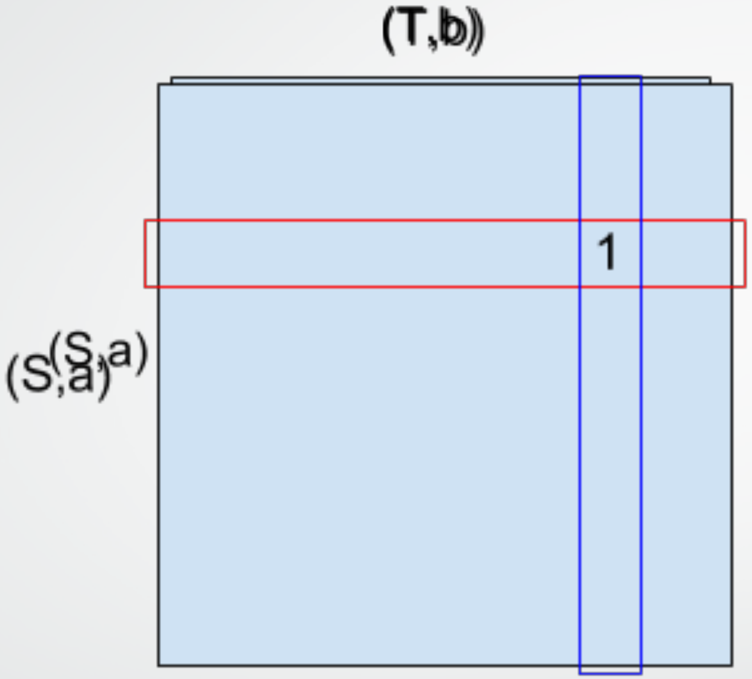
Quasi Poly Upper Bound [LMM'03] Hidden Clique Lower Bound [HK'09]

Roadmap



WHITE BOX REDUCTION !

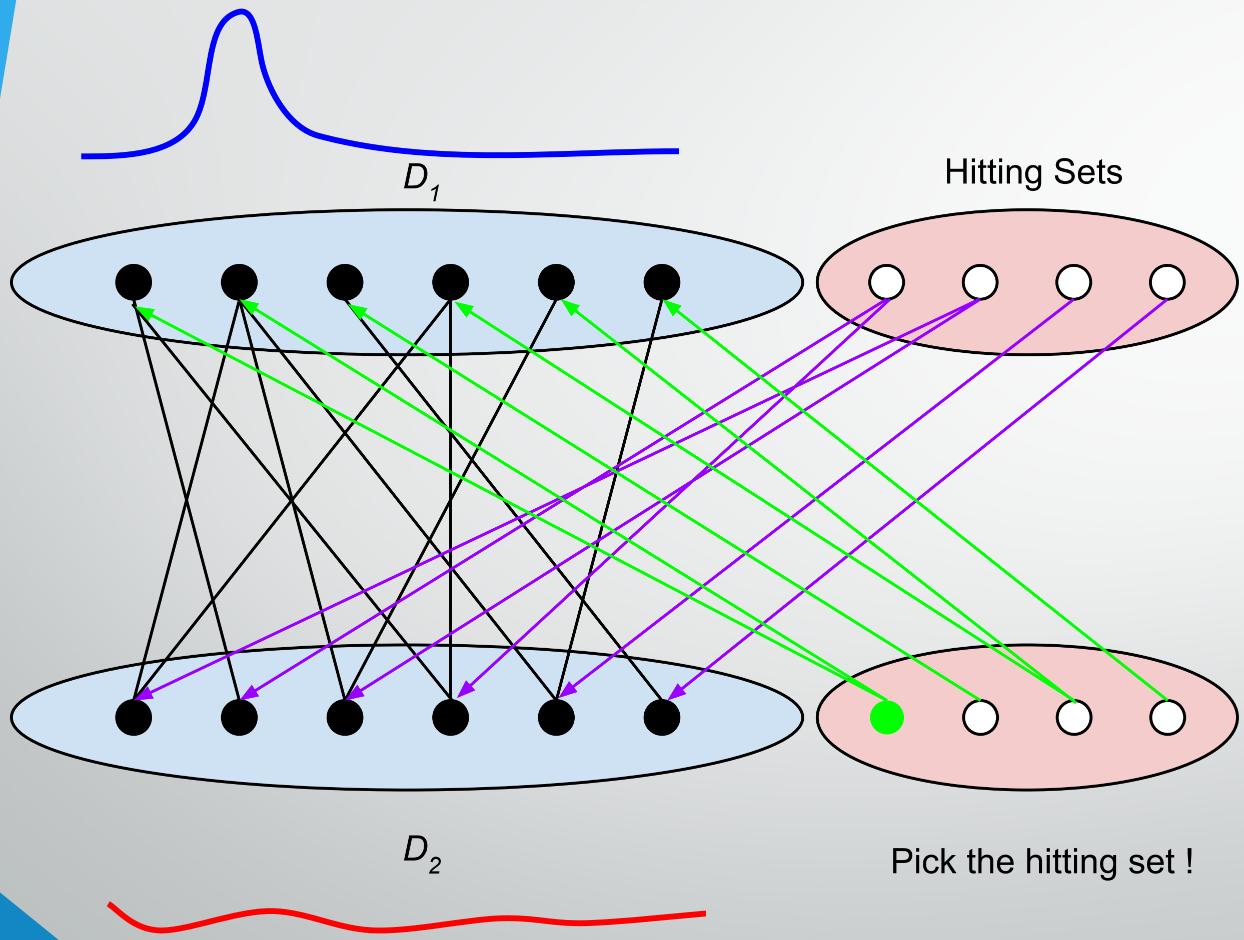
The "Trivial" Reduction



- Each row corresponds to Alice's choice of $O(\sqrt{n})$-size tuple of challenges and corresponding assignment to it (columns Bob's)
- Payoff 1 if it has "sufficient" number of constraints in between + satisfy all of them. Otherwise 0 (same for Alice and Bob)
- Problem with the reduction ?

    Distribution over the challenges can be messed up by the players (NO REFEREE !)

Hitting Set Argument in 2-player game



- Inspired by [HK'09]
- Consider event space on the strategy space
- Punish the opponent if he puts too much mass on particular event !
- Uniformity over the events enforced !
- Any "Good" approximate Nash must have "almost" uniform marginals among the hitting Set Events!
- **Goal) Construct a good "Hitting Set"**

New Payoff Matrix

Original Game	Hitting Set (zero-sum)
Hitting Set (Zero-sum)	0

The Hitting Set?

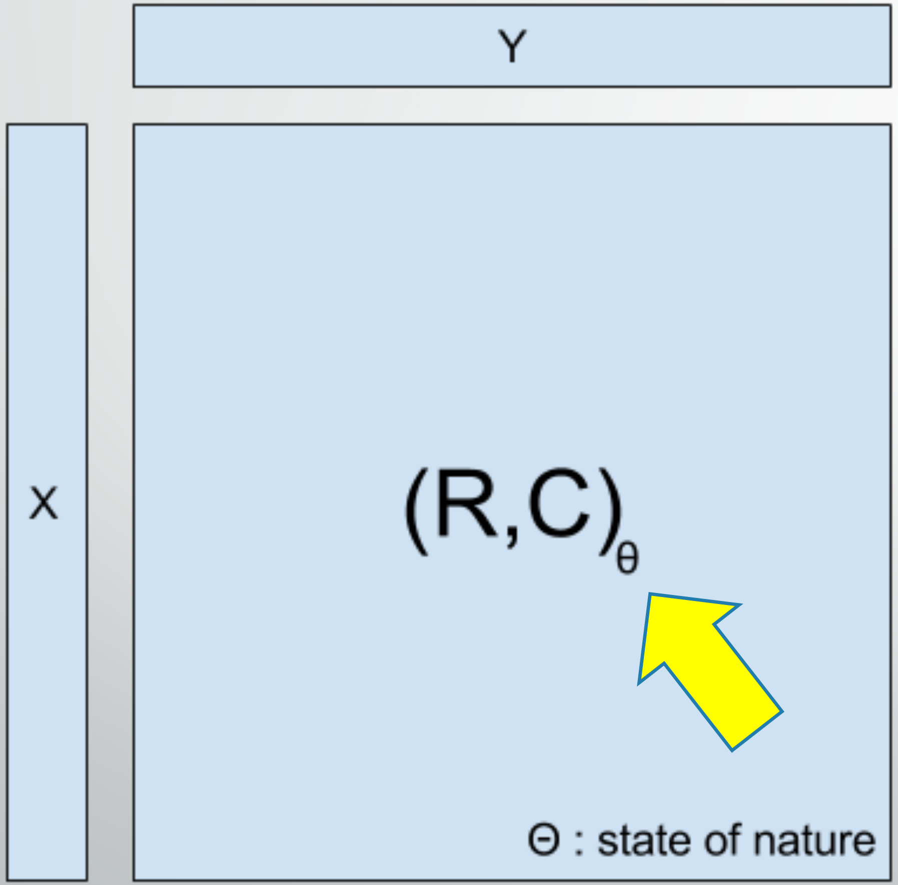
- **Appearance of x_i!** Bundle these events together to create the hitting set !
- If Alice (or Bob) focus on particular challenge, the opponent can punish (and him/herself can gain)!

➢ Uniform Marginals over the singletons (but unfortunately not over the sets)

➢ This suffices due to "**freeness**" of the distribution

Table of Contents

Extension:
Signaling in zero-sum **Bayesian Game**



- Bayesian Game ?
- State of Nature included (game changes depending on some unknown randomness)
- Players have limited knowledge (in terms of the payoff) of the game
- Zero-Sum game → Easy for compute!

Signaling in zero-sum Bayesian Game

- Given an instance of **zero-sum** Bayesian game over $\theta \in \Theta$ equipped with distribution λ, find "convex decomposition" of the distribution, which maximizes Player 1's payoff via this decomposition!

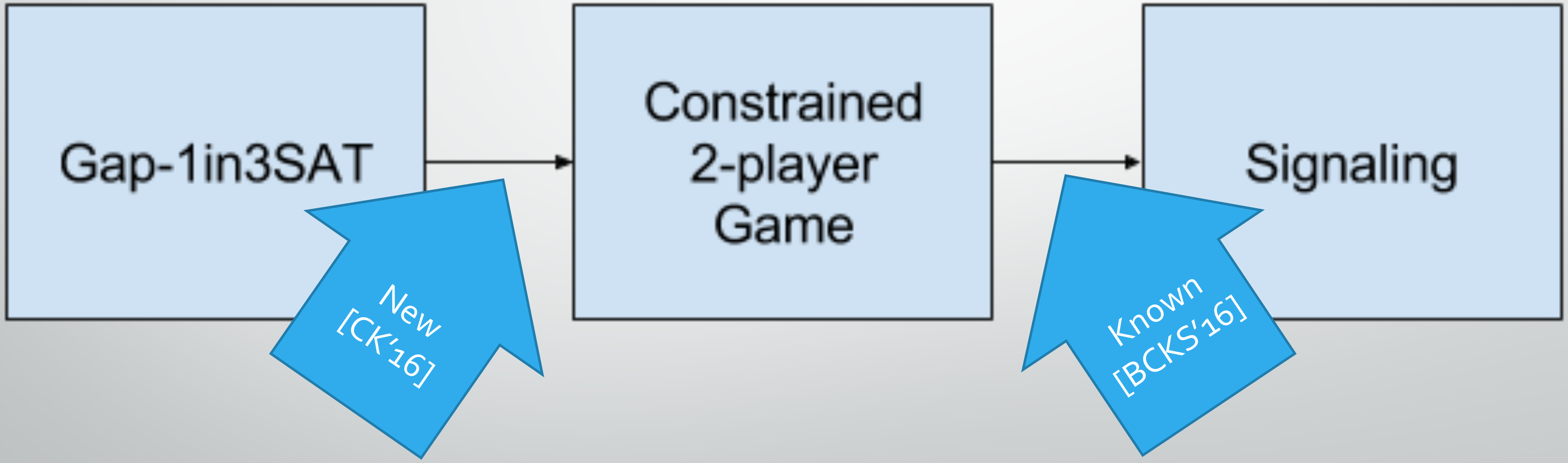
Easy

- Example) Road Network – what is the "optimal" signaling scheme
  - Full information isn't necessary optimal !
  - Market for Lemons !
- Quasi-polytime Approximation Algorithm [CCD+'15]
- Hardness based on Planted Clique Conjecture (search, decision) [Dug'15 , BCKS'16]
- Connection to Constrained 2-Player Game [BCKS'16]

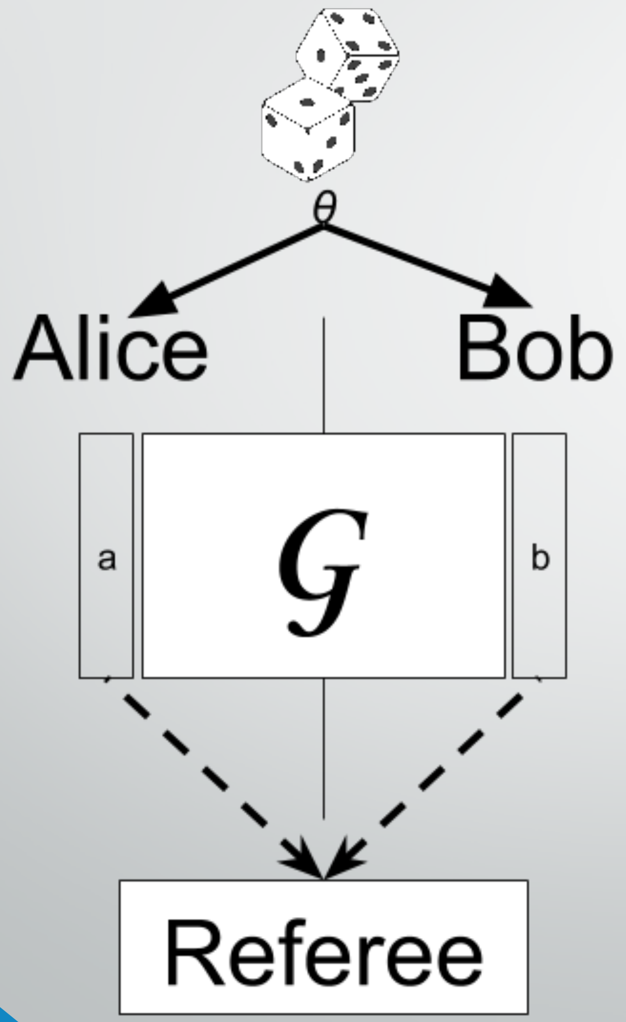
$$\sum_{\sigma} \lambda_\sigma P(\lambda_\sigma) \geq P(\lambda)$$

Such that $\sum_\sigma \lambda_\sigma = \lambda$

Roadmap


Gap-1in3SAT
Constrained 2-player Game
Signaling
New [CK'16]
Known [BCKS'16]

Constrained 2-player Game



- Same as before, but with additional guarantee in completeness
  - Let (a_θ, b_θ) be the set of good ϵ-NE. Then there exists some distribution λ such that $E_{\theta\sim\lambda}[\mathrm{b}_\theta]$ is "uniform"
- Arguing about the convex hull of good ANE !
- NOT A BAYESIAN GAME !
- "Constrained" in a sense that Bob must provide <u>marginally uniform strategy</u>!

Technical Challenges …

- Overall marginal of Bob’s strategy to be “**uniform”**
- The original reduction fails due to limited good assignment on 2CSPs
  i.e. it’s not uniform over the challenges and alphabets !
- Add all ”possible” **permutation** over the challenges which Alice gets to pick !
- Additional cheating strategy with the permutations …
  - Permutation adds more ways to “cheat”
- Impose additional Hitting Set Events to punish such strategies.

Can talk more about the precise Hitting Set Offline

Summary

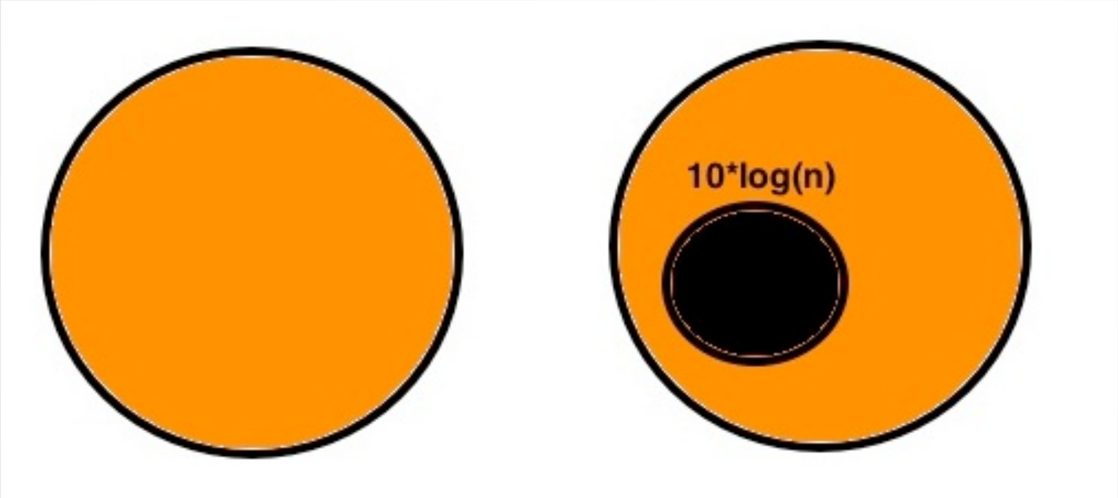
- Hitting Set Technique !
- Transform 2 Prover “Game” to strategic 2-Player “Game” with very little change in value with sub-exponential blowup
- Well-spread condition can be imposed in general
- Any other applications ???

Application: Entropy Argument

Based on [Braverman **K** Rubinstein Weinstein ' 15]

Table of Contents

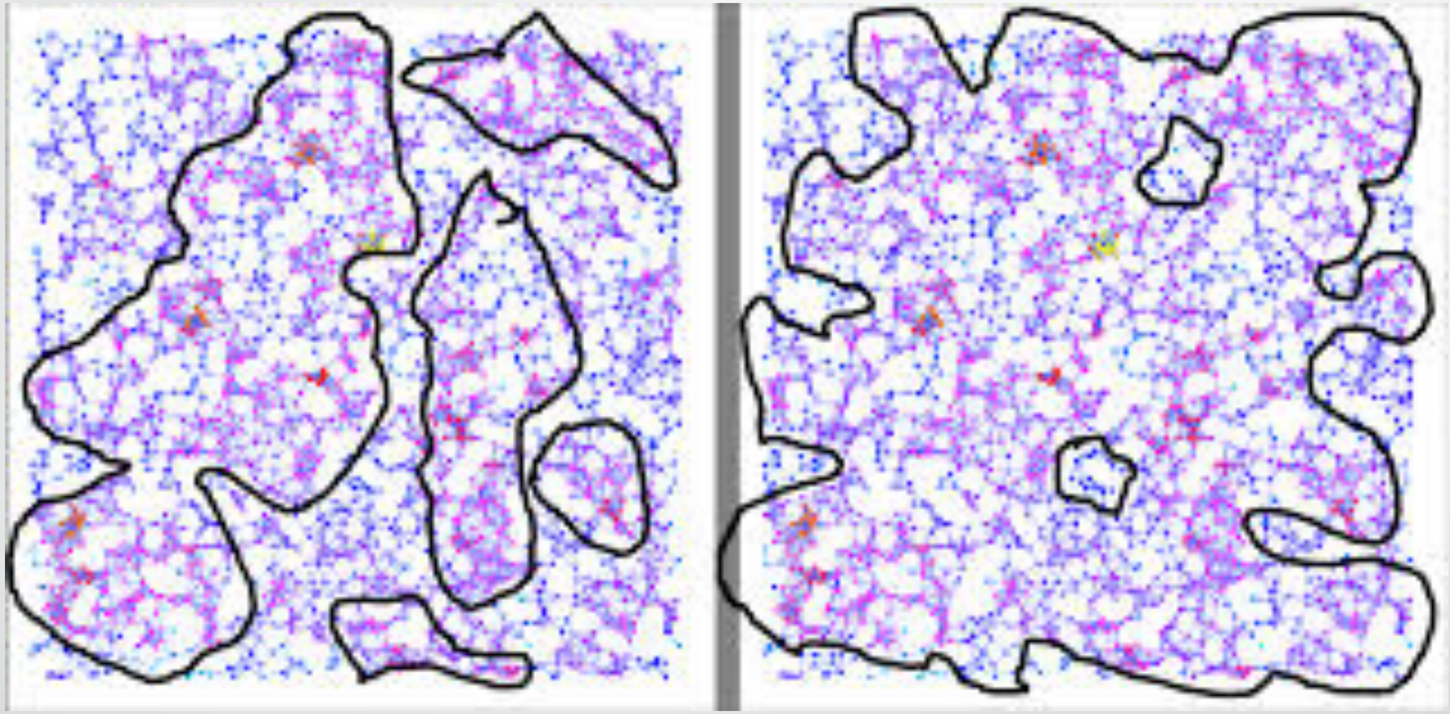
Reminder: Planted Clique Problem



- Distinguish between the following two cases :
  - $G\left(n,\frac{1}{2}\right)$ vs. $G\left(n,\frac{1}{2}\right)$ planted with a clique of size $10\log n$
- Search Version : find such planted clique ?
- Is Average case crucial ?? Lower-Bound ??
- Quasi-polytime algorithm obvious!

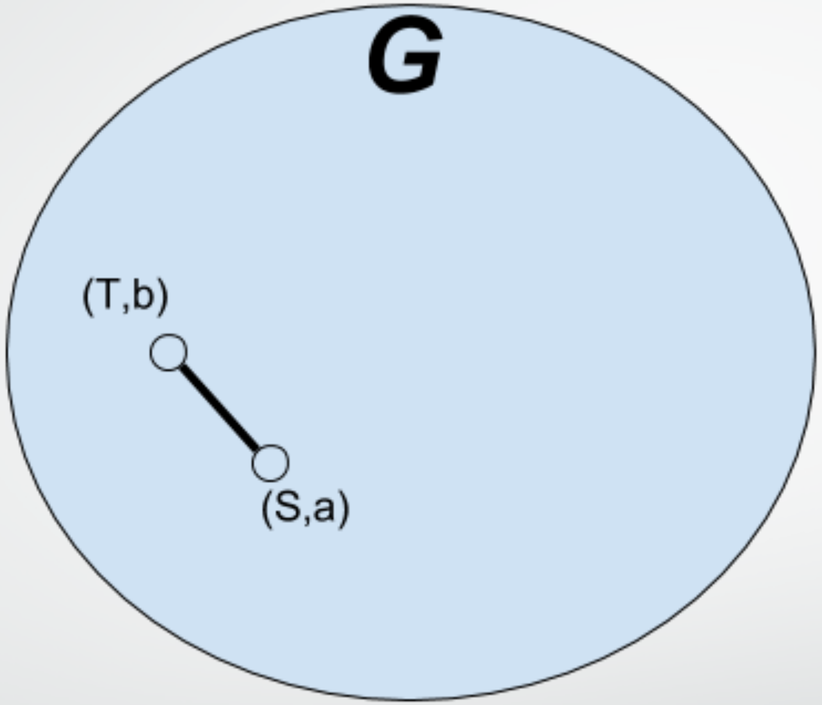
Spectral Based poly time algorithm by Alon & Sudakov only works for $\sqrt{n}$-size clique

Densest-k-Subgraph Problem



- Given $G = (V, E)$ and $k < |V|$, find $S \subset V$ of size k with the highest density
- Worst case version of Planted Clique Problem
- Quasi-polynomial "additive" approximation is known [FS'97, Bar'15]
- Perfect Completeness : Clique vs. a bit less than a clique

Very Simple Reduction



- Each vertex represents a partial assignment to $\tilde{O}(\sqrt{n})$ set of challenges
- $(u, v) \in E$ if u and v has no conflicts in the assignment
- Birthday Repetition => Most of the time, there is a checking between (u, v)

(Modification of FGLSS-graph : the reduction SAT to MAX-Clique)

Yet Intricate Analysis

Entropy Argument
(Inspired by Parallel Repetition)

- Main Observation : Define labeling on the vertices. Consider the distribution on the labeling for k-sized subset S. Then the entropy of the distribution is $\log k$
- Sum of the entropy from the assignment (a) + entropy from the choice of variables (S) must be $\log k$ as well
- If the entropy small from the choice of variables → Conflicts due to multiple choice of vertices from same set !
- If the entropy small from the choice of assignments → value of the original game kicks in !

Open Problems

- Hardness of finding stable community?
  - Definitely needs a new idea due to the instability condition on the solution
- Amplification of Densest k-subgraph completeness vs. soundness
  - A lot of evidences to believe so …
- Connection with un-ordered parallel repetition ? UG ? **SSE** ?
- Problems with Quasi-poly running time algorithms ? (Graph Isomorphism?)
- Finding a good open problem might be an open problem

Any Questions?