Sampling in Large Matrices and Approximation of Max-CSP

Marek Karpinski,
Univ. Bonn.

SAMPLING IN LARGE MATRICES AND APPROXIMATION OF MAX-CSP.

MAREK KARPINSKI,
UNIV. BONN.

(JOINT WORK WITH N. ALON, F. DE LA VEGA, AND R. KANNAN)

PROBABILISTIC SETTING:

GIVEN A MATRIX

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & \cdots & & \\ \vdots & & & \vdots \\ & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix},$$

n VERY LARGE

PROBABILISTIC SETTING:

GIVEN A MATRIX

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \cdots & & \vdots \\ \vdots & & \boxed{S} & \vdots \\ & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix},$$

n VERY LARGE

WHAT CAN WE LEARN ABOUT A FROM A RANDOM SM S?

PROBABILISTIC SETTING:

GIVEN A MATRIX

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \ddots & & \vdots \\ & \boxed{S} & & \\ & & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix},$$

n VERY LARGE

WHAT CAN WE LEARN ABOUT A FROM A RANDOM SM S?

RANDOM SUBMATRIX S OF A FIXED SIZE (UNIFORM SAMPLING; "COINS" HAVE TO BE TOSSED "BLINDLY" BEFORE A IS READ).

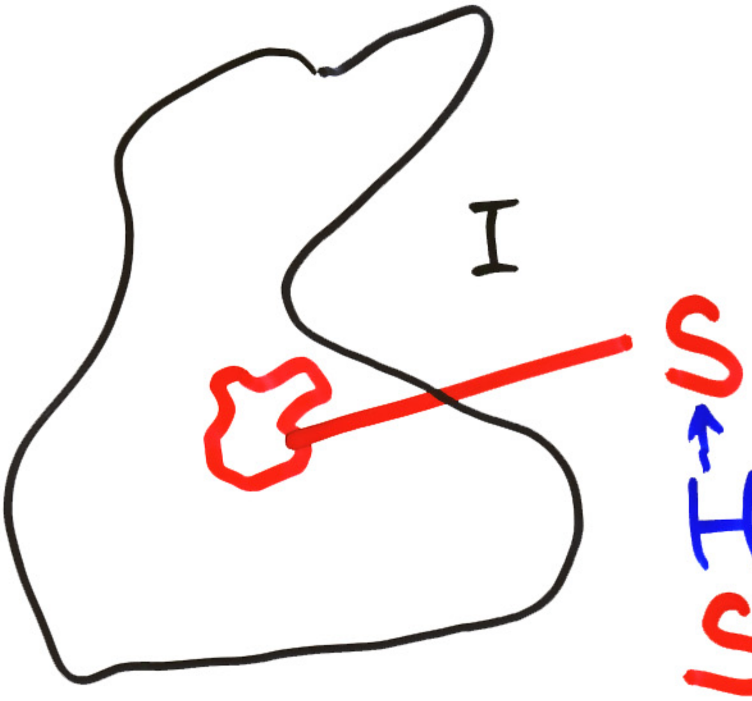
IF A IS TO DESCRIBE AN OPTIMIZATION PROBLEM INSTANCE (LIKE GRAPH PARTITIONING), WHAT IS A RELATION BETWEEN THE OPTIMA OF A AND S?

BASIC ISSUE IN COMBINATORIAL OPTIMIZATION AND PROPERTY TESTING.

COMPUTATIONALLY VERY DIFFICULT, IN MOST CASES NP-HARD.

OCCURS NATURALLY IN MASSIVE DATA PROBLEMS, WHERE ONLY A SMALL SAMPLE IS STORABLE IN RAM-MEMORY

(→ ALGORITHM SHOULD PROCESS SAMPLE ONLY AND GIVE GOOD ESTIMATE FOR THE WHOLE PROBLEM)

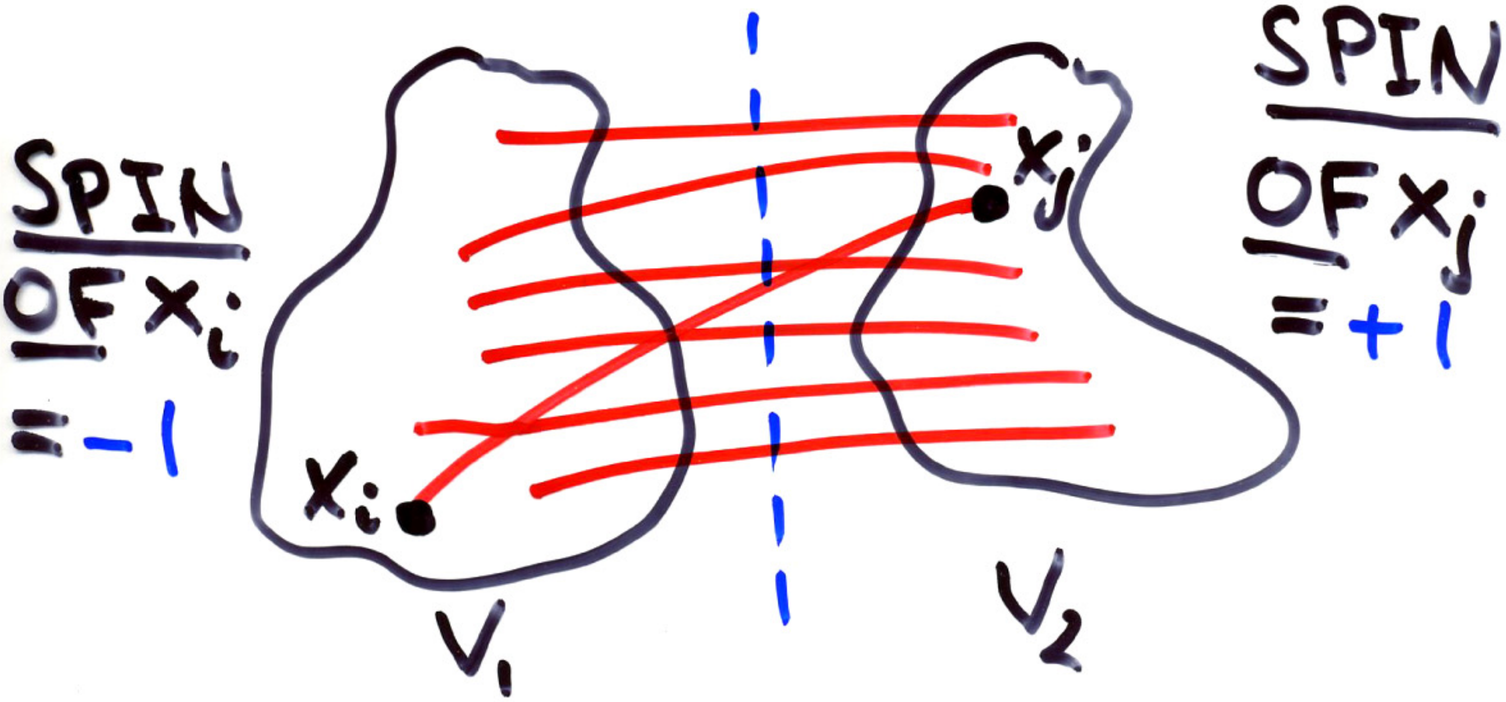


HARD CORE
SAMPLE

S **VERY** TINY
(CONSTANT **SIZE**?),
DOES **OPT**. OF
S **TELLS** US
"**SOMETHING**"
ABOUT **OPT**. OF I
?

MAX-CUT:

GIVEN A GRAPH $G=(V,E)$. PARTITION V INTO TWO GROUPS SO AS TO MAXIM. THE # OF EDGES BETWEEN THEM.



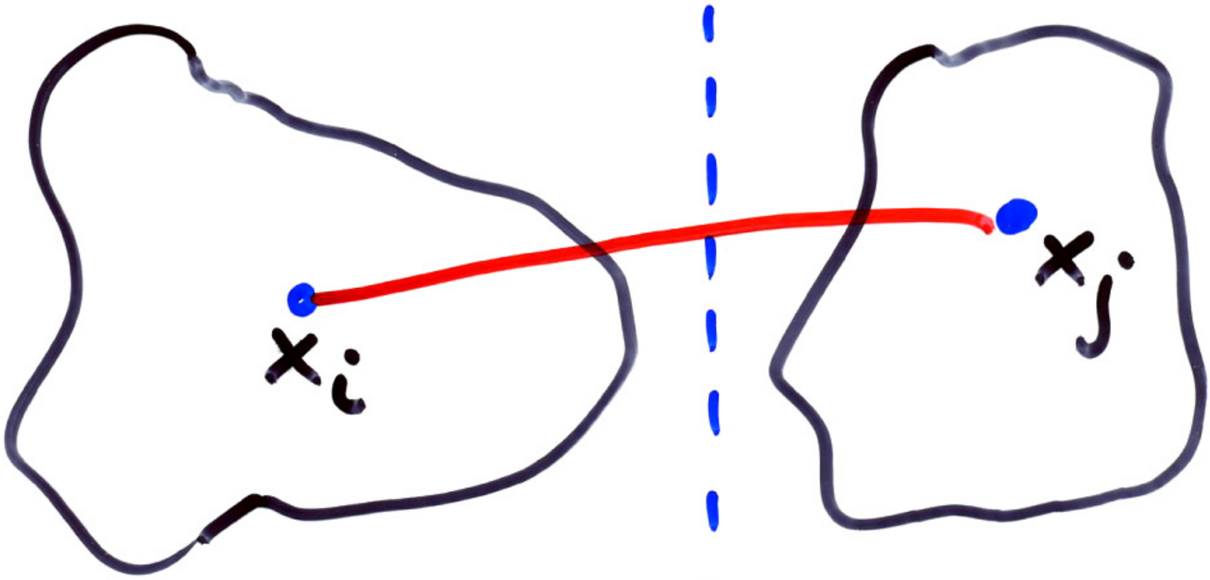
MAX-2CSP

INTERPRETATION:

$$\boxed{G}$$

$$\downarrow$$

$$\mathcal{C} = \{ x_i \oplus x_j = 1 \mid \{i,j\} \in E \}$$



$x_i = \underline{\underline{0}}$ $\qquad$ $x_j = \underline{\underline{1}}$

$$\boxed{x_i \oplus x_j = 1}$$

MAX/MIN r-CSP

EQUIVALENT TO THE FOLLOWING:

GIVEN A SET OF r-ARY BOOLEAN FNCT.'s $f_1, f_2, \ldots, f_m$. CONSTRUCT AN ASSIGNMENT $x \in \{0,1\}^n$ SO AS TO MAXIMIZE/MINIMIZE THE NUMBER OF SATISFIED FNCT.'s f_i.

$\Downarrow$

$$\boxed{\underline{\text{MAX-CUT}} \in \underline{\text{MAX-2CSP}}}$$

NP-HARD

IN EXACT

SETTINGS. ✓

APPROX.

COMPLEXITY?

Two Notions of Approx.'s:

- **Absolute:**

$$|OPT - Y| \leq \underbrace{r}$$

- **Relative:**

$$\text{Max}\left\{ \frac{Y}{OPT}, \frac{OPT}{Y} \right\} \leq \underbrace{r}$$

[Y = Costs of Solution S.]

$\mathcal{A}$ - **PROBABILISTIC**

$$\Pr\left[\mathcal{A} \text{ ``}\underline{\text{APPROX}}\text{''} \ P\right] \geq 1 - \delta,$$

$$\underbrace{\delta < 10^{-100}}.$$

$\mathcal{A}$ IS A PTAS
IF $\mathcal{A}$ APPROX.
WITH $\underbrace{r = 1+\varepsilon}$
(FOR ALL $\varepsilon > 0$).

RUNNING TIME
EFFICIENCY OF $\mathcal{A}$:

$$\boxed{|I|^{O(1)}}$$

DEPENDING ON FIXED $\varepsilon > 0$.

ϱ-ABSOLUTE PTAS (ϱ-APTAS)

CONDITION:

$|OPT - y| \leq \varepsilon\varrho$,

FOR ALL $\varepsilon > 0$.

IF OPT IS OF ORDER ϱ ($OPT = \Theta(\varrho)$) ($\rightarrow$ DENSE INSTANCES), EXIST. OF ϱ-APTAS $\Longrightarrow$ EXIST. OF PTAS)

IN OUR CASE THE KERNEL ϱ HAVE TO BE EQUAL TO n^2, AND THE ISSUE IS THE EXISTENCE OF n^2-APTAS & COMPUTING Y, S.T. $|OPT - Y| \leq \varepsilon n^2$ FOR ALL $\varepsilon > 0$.

Polynomial Matrix Repres. of Max-Cut:

$$P = \sum_{\{i,j\}\in E} \left(x_i(1-x_j) + x_j(1-x_i) \right),$$

$$x_i, x_j \in \{0,1\};$$

$$P = \sum a_{ij} x_i x_j + \sum b_i x_i + d.$$

Polynomial Matrix Repres. of Max-Cut:

$$P = \sum_{\{i,j\}\in E} \left(x_i(1-x_j) + x_j(1-x_i) \right),$$

$$x_i, x_j \in \{0,1\};$$

$$P = \sum a_{ij} x_i x_j + \sum b_i x_i + d.$$

SPIP

REDUCES TO THE PROBLEM OF **MAXIMIZING** A HOMOGENEOUS DEG. 2 POLYNOMIAL OVER A BOOLEAN CUBE $C = \{0,1\}^N$ $(\{-1,1\}^N)$,

$$\max_{y \in C} \sum A_{ij} y_i y_j$$

MORE GENERALLY

REDUCES TO THE PROBLEM OF MAXIMIZING A HOMOGENEOUS DEG. ~~2~~ r POLYNOMIAL OVER A BOOLEAN CUBE $C=\{0,1\}^N$ $(\{-1,1\}^N)$,

$$\max_{y\in C} \sum \underbrace{A_{ij}}_{r\text{-DIM. ARRAY}} y_i y_j$$

(+) CUT-NORM

OF A APPROXIM.

$(\|A\|_c)$.

⊕ CUT-NORM OF A <u>APPROXIM.</u> ($\|A\|_c$).

→ IT "<u>HAPPENS</u>" THAT:

<u>OPT</u> $= \underbrace{\|A\|_c}$

NOTATION:

GIVEN FINITE **SETS** $V_1, V_2, \ldots, V_r$,

AN r-**DIMENSIONAL ARRAY** A ON $V_1, \ldots, V_r$ IS A **FUNCTION**

$$A : V_1 \times V_2 \times \ldots V_r \to \mathbb{R}$$

($A(i_1, i_2, \ldots, i_r)$ IS AN **ENTRY** OF A).

FROBENIUS NORM OF A :

$\|A\|_F$ = SQUARE ROOT OF THE SUM OF SQUARES OF ALL ENTRIES.

LET $S_1 \subseteq V_1, S_2 \subseteq V_2, \dots, S_r \subseteq V_r$, **DEFINE** THE **QANTITY**:

$$A(S_1, S_2, \dots, S_r) = \sum_{(i_1,\dots,i_r) \in S_1 \times S_2 \times \dots \times S_r} A(i_1, i_2, \dots, i_r).$$

A **CUT-NORM** OF **A**:

$$\|A\|_C = \max_{S_1 \subseteq V_1, \dots, S_r \subseteq V} |A(S_1, \dots, S_2)|$$

CUT-NORM RELATES FOR $r=2$ - TO GROTHENDIECK NORM, AND SD-PROGRAMMING,

$$\|A\|_G = \max_{\substack{x_i, y_j \in \\ \{-1,1\}}} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} x_i y_j \right\}$$

$$\boxed{\|A\|_C \leq \|A\|_G \leq 4\|A\|_C}$$

EXAMPLE:

$r=2$, A IS A MATRIX WITH A SET OF ROWS INDEXED BY E' AND A SET OF COLUMNS INDEXED BY V, FOR A GIVEN GRAPH $G=(V,E)$,

$V=\{v_1, v_2, \ldots, v_n\}$,

$E=\{e_1, e_2, \ldots, e_m\}$, AND

$E'=\{e_1, \ldots, e_m, e'_1, \ldots, e'_m\}$.

$$A = \begin{array}{cc} & V \\ E' & \begin{bmatrix} & \vdots & \\ \cdots & a_{ki} & \cdots \\ & \vdots & \end{bmatrix} \end{array}$$

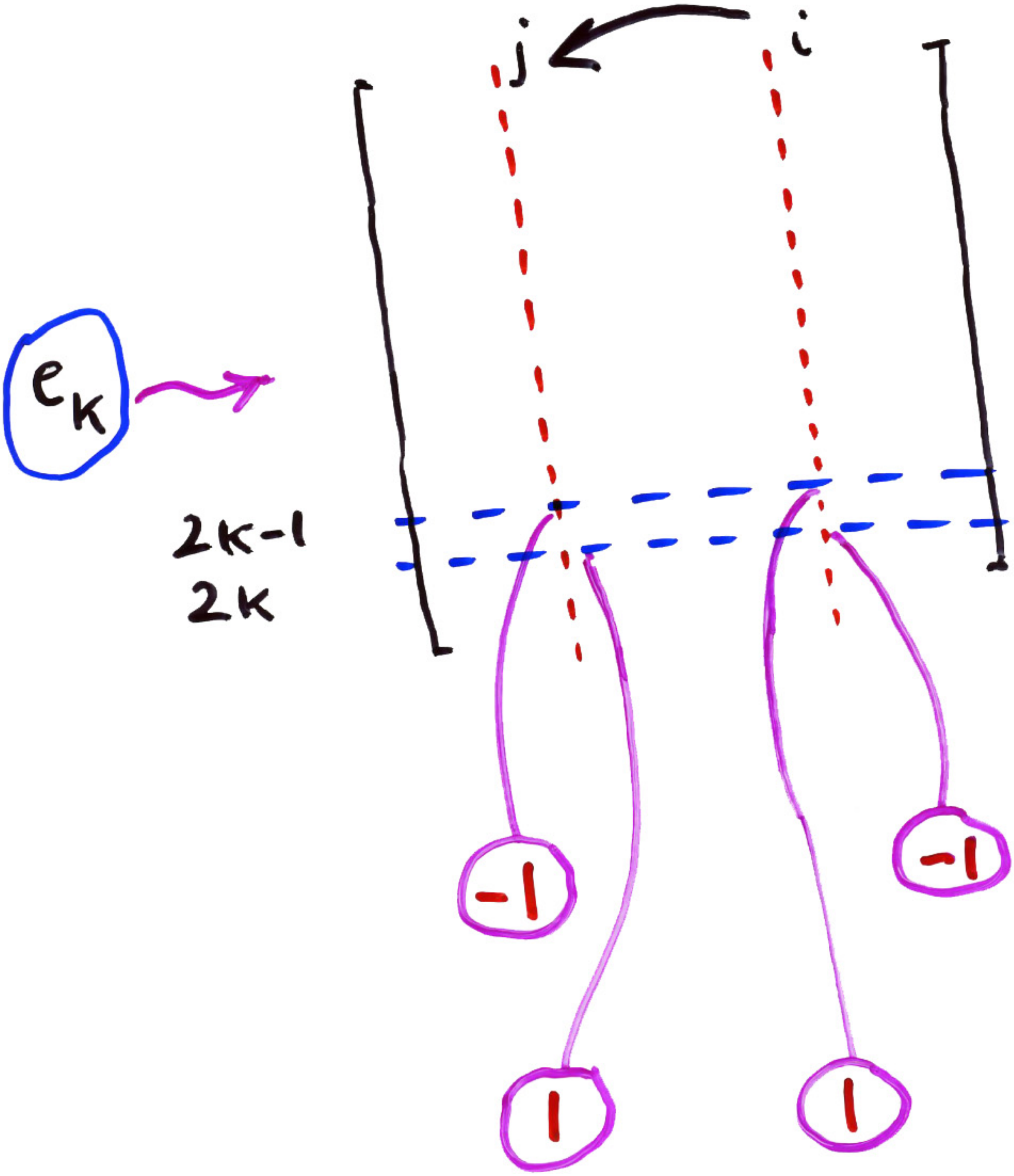
ORIENT $G=(V,E)$ ARBITRARILY, E.G.

$v_i \longrightarrow v_j$ IFF $i>j$.

LET $e_k = (v_i \longrightarrow v_j)$,

SET: $\left[\begin{array}{l} a_{2k-1,i} = \\ \quad a_{2k,j} = 1 \end{array} \right]$;

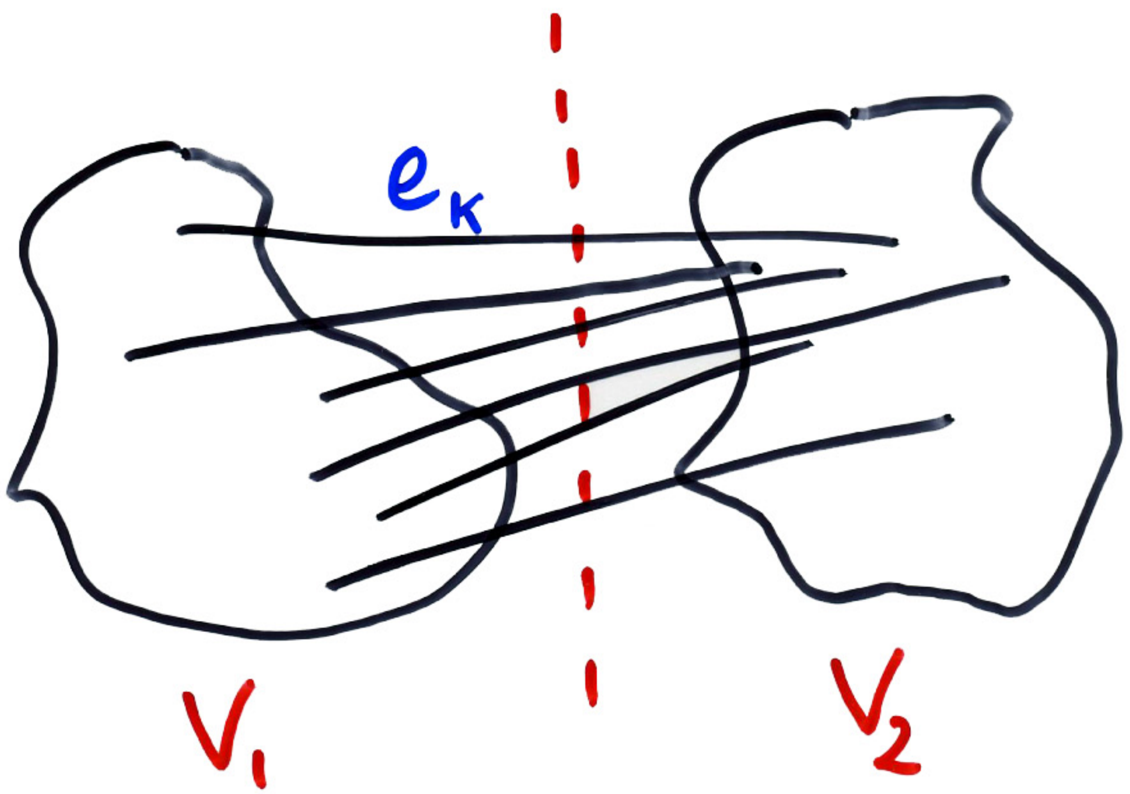
AND $[a_{2k-1,j} = a_{2k,i} = -1]$.



THE REST OF ENTRIES ARE ALL SET TO 0.

ANY CUT OF

G :



INDUCES A SET OF EDGES e_k IN A CUT, AND THE SET OF 1-ENTRIES IN A. THUS,

$$\|A\|_c = \text{MAX-CUT}(G).$$

MAX-rCSP CAN BE REDUCED TO THE PROBLEMS OF MAXIMIZING POLYNOMIALS OF DEGREE r OVER THE BOOLEAN CUBE (AS WE DID USING AN SPIP FOR MAX-CUT PROBLEM), AND COMPUT. OF $\|A\|_c$ FOR r-DIM. ARRAYS A.

APPROXIMATION OF $\|A\|_C$.

ON A RANDOM SUBSET OF SIZE $\tilde{\Theta}\left(\log(1/\varepsilon)/\varepsilon^4\right)$ $(= \underbrace{q})$.

ASSUMPTIONS (*) ON A:

$\|A\|_C \leq \varepsilon n^r$, $\|A\|_\infty \leq \frac{1}{\varepsilon} B(r)$,

$\|A\|_F \leq 2^{2^r} n^{r/2}$.

THEN, A RANDOM INDUCED SUBARRAY H OF A SATISFIES

$$\|H\|_c \leq \underbrace{C(r) \cdot q^{r}}$$

W.H.P.

THE OTHER DIRECTION IS EASY:

IF $\|A\|_c$ IS HIGH, THEN SO IS $\|H\|_c$.

$\boxed{\text{W.H.P.:}}$

MORE EXPLICITLY:

FOR

$$q \geq 10^6 r^{12} \frac{1}{\delta^5 \varepsilon^4} \operatorname{LOG}\left(\frac{4}{\varepsilon^2}\right),$$

$\delta, \varepsilon > 0,$

$$\|H\|_c \leq 2^{2^r + q} \frac{\varepsilon}{\sqrt{\delta}} q^r$$

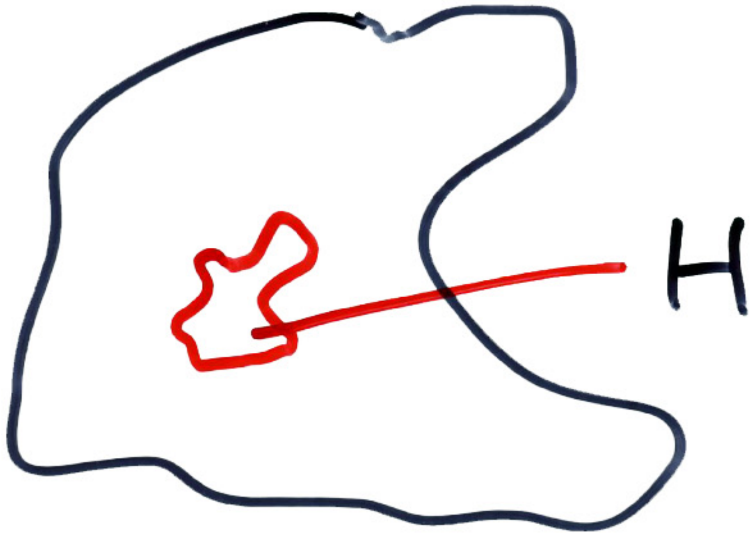
WITH PROB. $\geq 1 - \delta$.

THE COMPUTATION OF $\|H\|_c$ (WE NEED ONLY ABS. APPROX.) ON A SMALL SAMPLE CAN BE DONE BY KNOWN METHOD OF CUT-ARRAY DECOMPOSITION.

RESULTING TIME:

$$2^{O(1/\varepsilon^2)}.$$

SO **FAR**, WE WERE **ABLE** TO DEAL WITH "DOWN-STAIRS" **SITATION** FROM THE **WHOLE** ARRAY TO THE **SMALL** RANDOM SUBARRAY H.



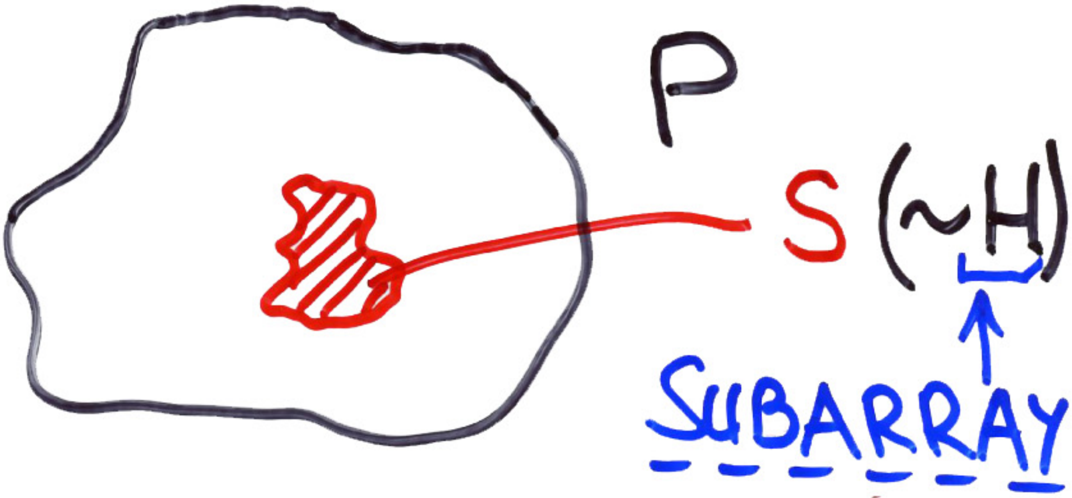
HOW **ABOUT** REVERSED ("UP-STAIRS") **DIRECTION**?

TO **RELATE** $\|H\|_c$ **WITH** THE $\|A\|_c$ (**OPT** OF AN **INSTANCE**).

LP Part:

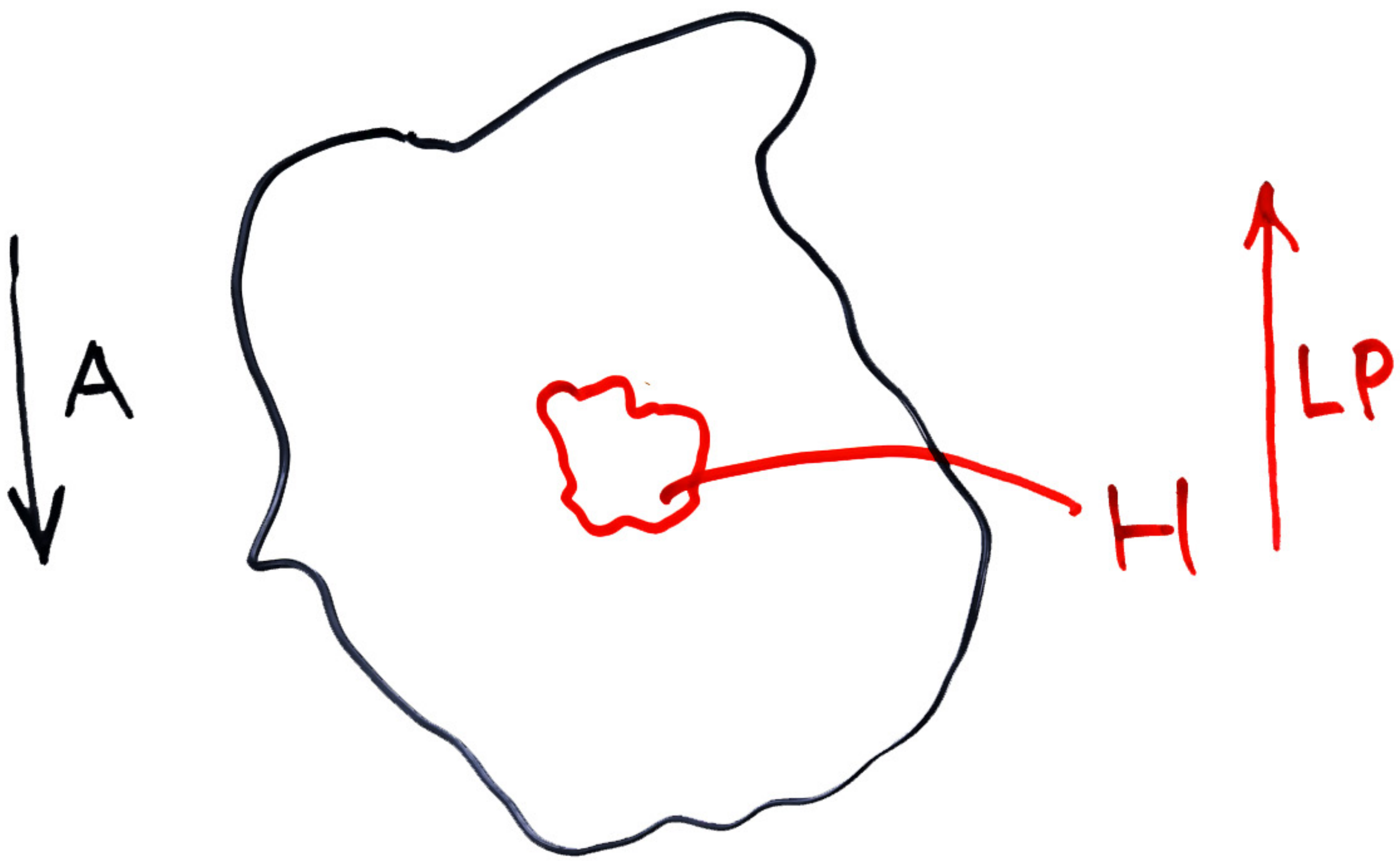
Take a linearized SPIP for an inst. F which is an ILP, call it P.

$|S| = \Theta(1/\varepsilon^4)$



If $\alpha \geq OPT_P$, then
$\alpha' \geq OPT_{P^S}$.

A SITUATION:



$$\beta \leq \|A\|_c \leq \varepsilon n^r$$

$$\Downarrow \qquad \uparrow \text{LP}$$

$$\beta' \leq \underbrace{\|H\|_c \leq \alpha'}$$

[ILP P GIVES

NORMAL. FACTOR $\frac{n^r}{q^r}$

MAIN RESULT:

LET F BE AN INST. OF MAX-rCSP WITH n VAR.'S. FOR A RANDOM SAMPLE Q OF THE SET OF VAR.'S $\{x_1, \dots, x_n\}$ LET F^Q BE A RANDOM SUB-INST. INDUCED BY Q.

MAIN RESULT:

$$\left| \frac{n^r}{q^r} \mathrm{OPT}_{F^Q} - \mathrm{OPT}_F \right| \leq \varepsilon n^r$$

FOR $|Q| = q = O(\log(1/\varepsilon)/\varepsilon^4)$

LET F BE AN INST. OF MAX-rCSP WITH n VAR.'S. FOR A RANDOM SAMPLE Q OF THE SET OF VAR.'S $\{x_1, \dots, x_n\}$ LET F^Q BE A RANDOM SUB-INST. INDUCED BY Q. W.H.P.

MAIN RESULT:

$$\left| \frac{n^r}{q^r} \mathrm{OPT}_{F^Q} - \mathrm{OPT}_F \right| \leq \varepsilon n^r$$

FOR $|Q| = q = O(\log(1/\varepsilon)/\varepsilon^4)$

VERY RECENT IMPROVEMET OF $\boxed{\text{A-PART}}$ TO HARD CORE SIZE $\tilde{\Theta}\left(1/\varepsilon^2\right)$ BY RUDELSON AND VERSHYNIN USING SOME NEW TECHNIQUES OF BOURGAIN AND TZAFRIRI.

FURTHER RESEARCH:

IMPROVING THE HARD CORE COMPLEXITY OF THE ALGORITHMS.

ARE THERE ANY "MYSTERIOUS" PROB. INTRACTABILITY BARRIERS FOR GETTING DOWN TO, SAY, $\tilde{\Theta}(1/\varepsilon^2)$ HARD CORE BOUNDS?

HOW ABOUT "SOFT ENVELOPE" (RAPIDLY MIXING MCMC) METHODS FOR THAT PROBLEM

?

■