Sept. 4, 2023

Something about log-supermodular distributions

Shuji Kijima (Shiga University)

1. Tutte polynomial

--- As an introduction of log-supermodular

Tutte polynomial -- as an introduction of log-supemodular

The ***Tutte polynomial*** of a graph $G = (V, E)$ is given by

$$T_G(x, y) := \sum_{A \in 2^E} (x-1)^{r(E)-r(A)} (y-1)^{|A|-r(A)}$$

for $x, y \in \mathbb{R}$ where $r(A) = \max\{\, |F| \mid F \subseteq A \text{ is a forest (cycle free)} \,\}$,
i.e., *rank function* of the graphic matroid (a.k.a. cycle matroid).

Tutte polynomial contains a lot of information on G:

$T_G(1,1) = $ #spanning trees of G

$T_G(2,1) = $ #forests of G

$T_G(1,2) = $ #spanning subgraphs of G

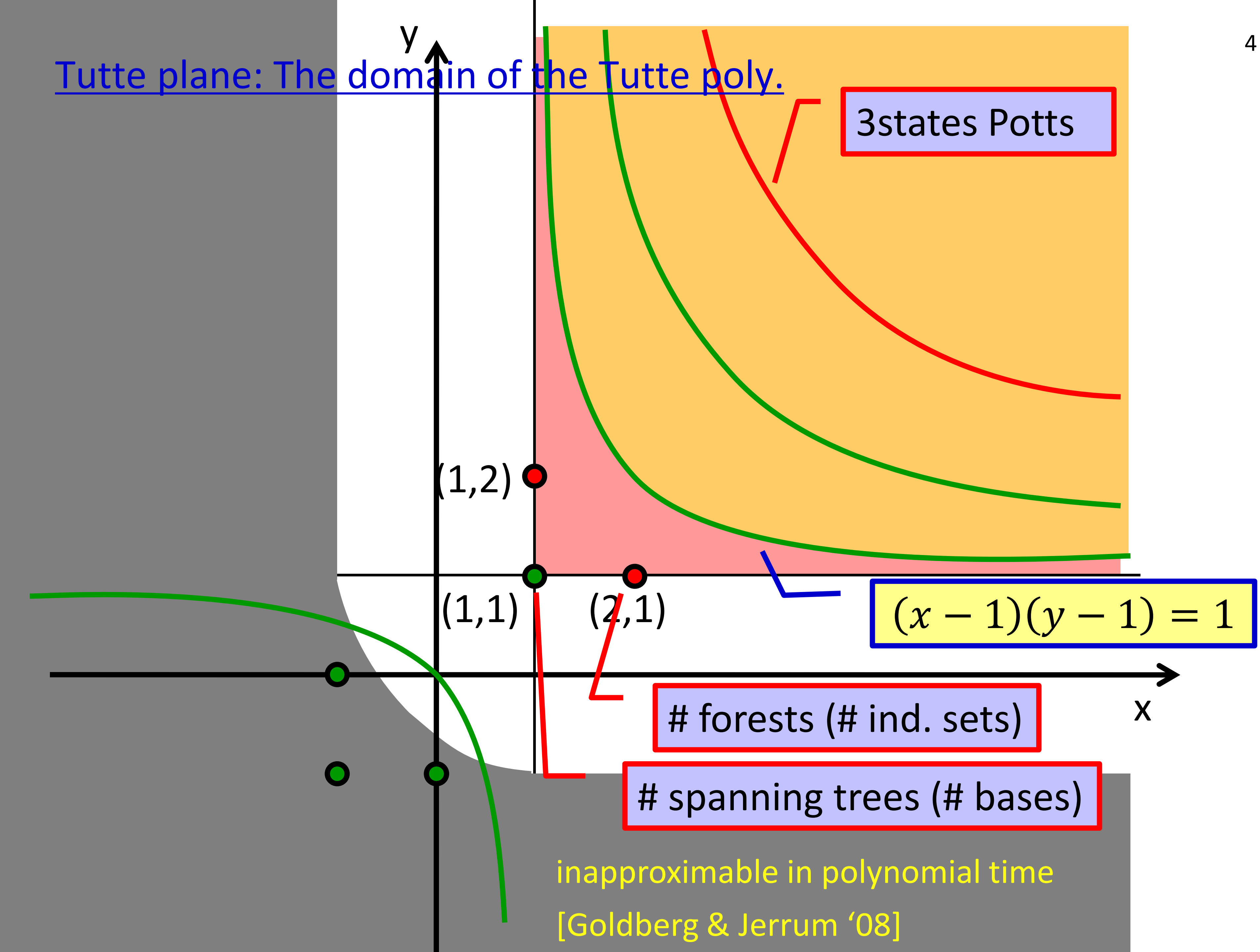
$H_q = \{(x, y) | (x-1)(y-1) = q\}$: part. func. Potts model w/ q-states

----

$T_G(x, 0)$: chromatic polynomial

$T_G(2,0) = $ #acyclic orientation

Tutte plane: The domain of the Tutte poly.



Tutte polynomial as a partition function

This term is positive for each $A \in 2^E$ if $x > 1$ and $y > 1$

$$T_G(x,y) := \sum_{A \in 2^E} (x-1)^{r(E)-r(A)} (y-1)^{|A|-r(A)}$$

Tutte poly. is regarded as the partition fnc. (normalizing const.) of …

Tutte polynomial as a partition function

This term is positive for each $A \in 2^E$ if $x > 1$ and $y > 1$

$$T_G(x,y) := \sum_{A \in 2^E} (x-1)^{r(E)-r(A)} (y-1)^{|A|-r(A)}$$

Tutte poly. is regarded as the partition fnc. (normalizing const.) of ...

a distribution π_G over 2^E, given by

$$\pi_G(X) := \frac{1}{T_G(x,y)} (x-1)^{r(E)-r(X)} (y-1)^{|X|-r(X)}$$

for $X \in 2^E$,

when $x > 1$ and $y > 1$.

Tutte polynomial as a partition function

This term is positive for each $A \in 2^E$ if $x > 1$ and $y > 1$

$$T_G(x,y) := \sum_{A \in 2^E} (x-1)^{r(E)-r(A)} (y-1)^{|A|-r(A)}$$

Tutte poly. is regarded as the partition fnc. (normalizing const.) of …

a distribution π_G over 2^E, given by

$$\pi_G(X) := \frac{1}{T_G(x,y)} (x-1)^{r(E)-r(X)} (y-1)^{|X|-r(X)}$$

for $X \in 2^E$,

when $x > 1$ and $y > 1$. Furthermore,

- ✓ π_G is log-supermodular if $(x-1)(y-1) \geq 1$,
- ✓ π_G is log-submodular if $(x-1)(y-1) \leq 1$

log-supermodular: $\pi_G(X)\pi_G(Y) \leq \pi_G(X \cup Y)\pi_G(X \cap Y)$

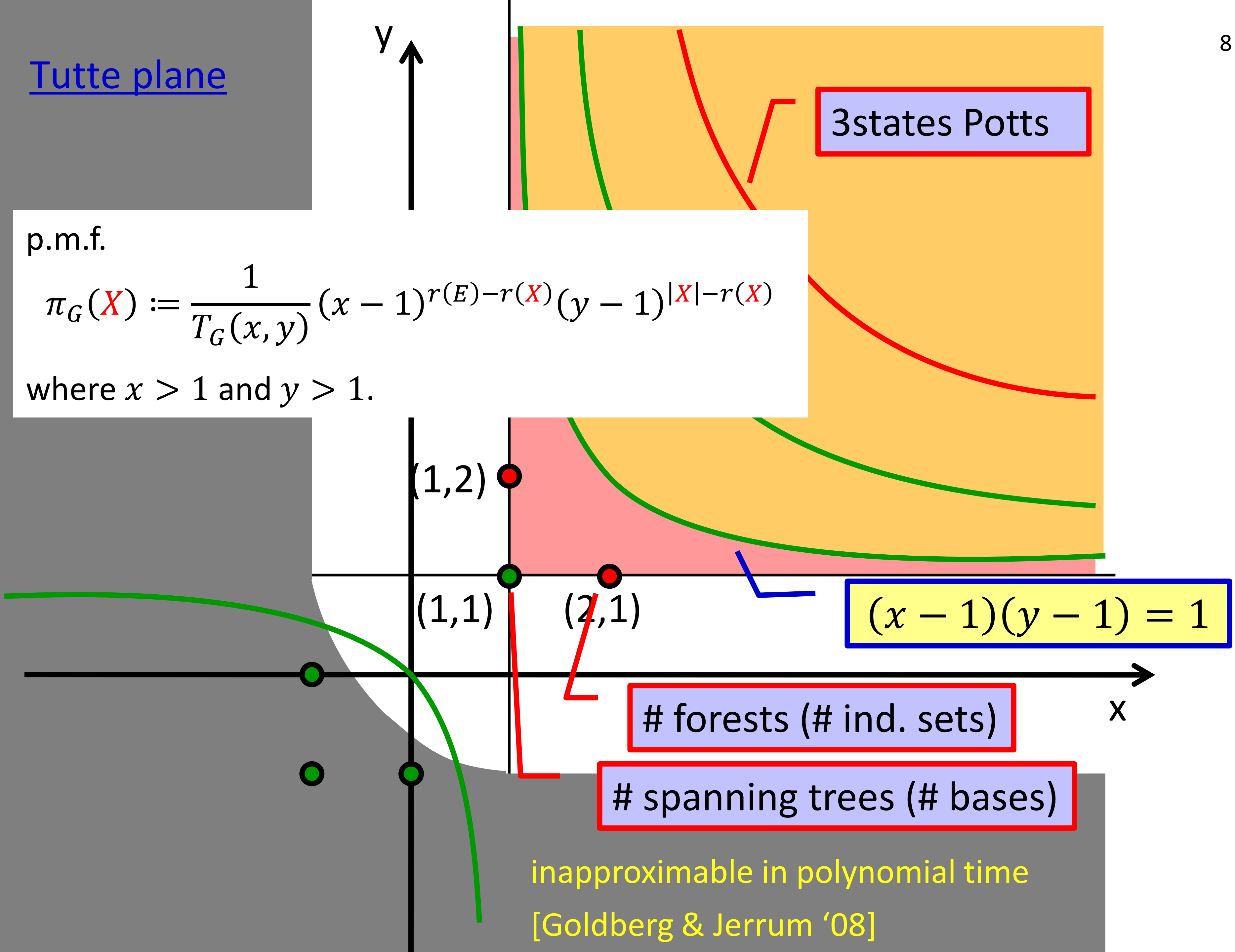
log-submodular: $\pi_G(X)\pi_G(Y) \geq \pi_G(X \cup Y)\pi_G(Y \cap Y)$

Tutte plane

p.m.f.

$$\pi_G(X) := \frac{1}{T_G(x,y)} (x-1)^{r(E)-r(X)} (y-1)^{|X|-r(X)}$$

where $x > 1$ and $y > 1$.



Goal of the talk

- Why log-supermodular?
  - ✓ Seemingly **"Tractable"**
    - ➢ Monotone coupling (cf. FKG ineq.), "log-concave" etc.
  - ✓ #BIS-**hard**
    - ➢ #Ideal, #stable matching

A challenge:
FPRAS or not

- What we (or I) know?
  - ✓ Log-concave?

    J. Nakashima, Y. Yamauchi, S. Kijima and M. Yamashita, Finding submodularity hidden in symmetric difference, SIAM Journal on Discrete Mathematics, 34:1 (2020), 571--585.
  - ✓ Subclass for #BIS-hard

    T. Fujii and S. Kijima, Every finite distributive lattice is isomorphic to the minimizer set of an M$^{\natural}$-concave set function, Operations Research Letters, 49:1 (January 2021), 1--4.

Goal of the talk

- Why log-supermodular?
  - ✓ Seemingly **“Tractable”**
    - ➢ Monotone coupling (cf. FKG ineq.), “log-concave” etc.
  - ✓ #BIS-**hard**
    - ➢ #Ideal, #stable matching

A challenge:
FPRAS or not

- What we (or I) know?
  - ✓ Log-concave?

    J. Nakashima, Y. Yamauchi, S. Kijima and M. Yamashita, Finding submodularity hidden in symmetric difference, SIAM Journal on Discrete Mathematics, 34:1 (2020), 571--585.

  - ✓ Subclass for #BIS-hard

    T. Fujii and S. Kijima, Every finite distributive lattice is isomorphic to the minimizer set of an $M^{\natural}$-concave set function, Operations Research Letters, 49:1 (January 2021), 1--4.

Tractable 1: Log-supermodular and monotone coupling

A set function $g: 2^N \to \mathbb{R}_{>0}$ is *log-supermodular* if
$$g(X)g(Y) \leq g(X \cup Y)g(X \cap Y)$$
holds for any $X, Y \in 2^N$, where $N = \{1,2,\dots,n\}$.

Define a transition from $X \in 2^N$ to $X' \in 2^N$ as follows

1. Choose $i \in N$ u.a.r.
2. Let $X' = \begin{cases} X \cup \{i\} & \text{w.p.} \frac{g(X\cup\{i\})}{g(X\cup\{i\})+g(X\setminus\{i\})}, \\ X \setminus \{i\} & \text{otherwise.} \end{cases}$

Prop. (cf. FKG ineq.)

The Markov chain admits a natural *monotone* coupling if (and only if) g is log-supermodular

Log-supermodularity is "iff condition" for a monotone CFTP

A naïve CFTP requires simulation from all the states (2^N, in our case).

If the Markov chain is ***stochastically monotone***, then two chains (from Max. and Min.) are sufficient for the CFTP algorithm.

J. G. Propp, D. B. Wilson, Exact sampling with coupled Markov chains and applications to statistical mechanics, Random Struct. Algorithms, 9(1-2), 223-252, 1996.

Thm. [K. @HJ '11]

A *reversible Hasse walk* on a distributive lattice has a *monotone update function*

⇔ its stationary distribution is *log-supermodular*

Remark.

We have an example that a hit-and-run chain (it's not a Hasse walk) admits a monotone CFTP for discretized Dirichlet distribution, which is not a log-supermodular distribution for some parameter [Matsui&K. `07]

Tractable 2: Log-supermodular vs. log concave

A set function $g: 2^N \to \mathbb{R}_{>0}$ is *log-supermodular* if
$$g(X)g(Y) \leq g(X \cup Y)g(X \cap Y)$$
holds for any $X, Y \in 2^N$, where $N = \{1,2,\ldots,n\}$.

- Equivalently, g is log-supermodular iff $-\log g$ is submodular, where a set function $f: 2^N \to \mathbb{R}$ is *submodular* if
$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y)$$
holds for any $X, Y \in 2^N$.
- Submodularity is often regarded as a discrete analogue of convexity:
  - ✓ f is submodular iff its Lovasz's extension is convex.
  - ✓ Minimization is in P, Maximization is NP-hard.

Set fncs.	Continuous fncs.
f is supermodular	f is concave
$\Leftrightarrow -f$ is submodular	$\Leftrightarrow -f$ is convex

- Log-supermodular is compared with log-concave:
  Maximum likelihood estimation is efficiently found.
- Log-submodular is compared with log-convex:
  Maximum likelihood estimation is hard in general.

Log-supermodular distributions

- Ferromagnetic Ising
- Tutte polynomial
- FKG inequality

Log-concave distributions

- Gaussian distribution

Q. Is there an efficient algorithm to sample from log-supermodular distribution?

Possible to sample from log-concave distribution, efficiently.

Goal of the talk

- Why log-supermodular?
  - ✓ Seemingly **“Tractable”**
    - ➢ Monotone coupling (cf. FKG ineq.), “log-concave” etc.
  - ✓ #BIS-**hard**
    - ➢ #Ideal, #stable matching

A challenge:
FPRAS or not

- What we (or I) know?
  - ✓ Log-concave?

    J. Nakashima, Y. Yamauchi, S. Kijima and M. Yamashita, Finding submodularity hidden in symmetric difference, SIAM Journal on Discrete Mathematics, 34:1 (2020), 571--585.

  - ✓ Subclass for #BIS-hard

    T. Fujii and S. Kijima, Every finite distributive lattice is isomorphic to the minimizer set of an $M^{\natural}$-concave set function, Operations Research Letters, 49:1 (January 2021), 1--4.

Intractable: What is #BIS? #BIS is a counting problm

Prob. #BIS

Given $G = (U, V; E)$ **B**ipartite graph.

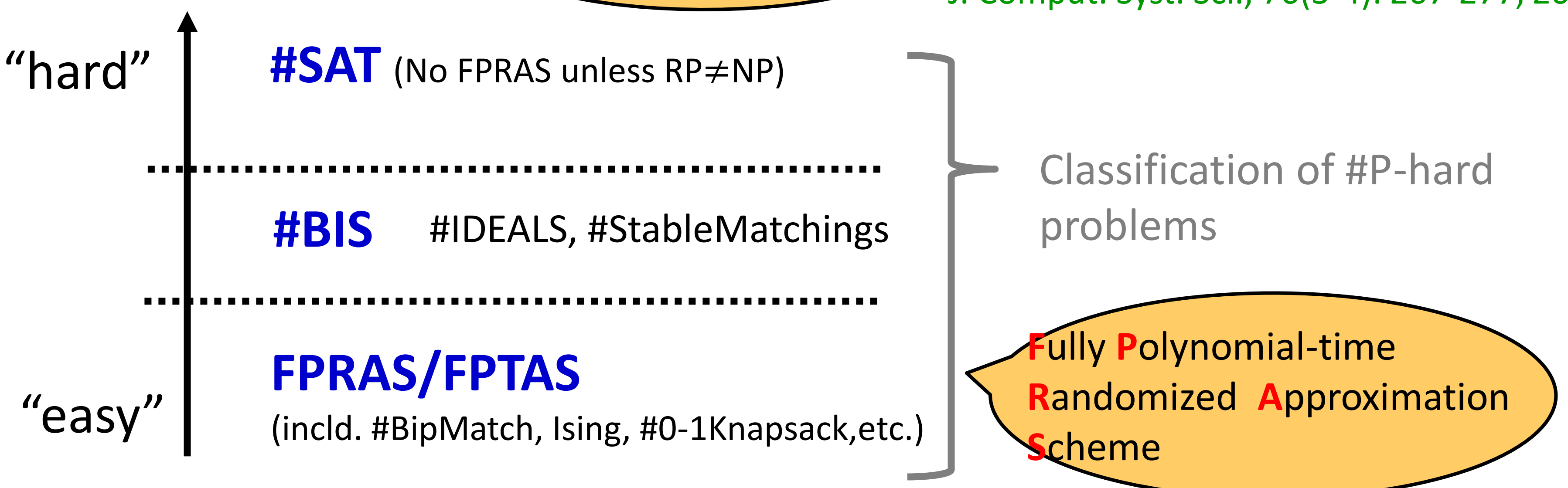
Count the number of **I**ndependent **S**ets, where

$X \subseteq U \cup V$ is an *independent set* if $\{x, y\} \notin E$ for any $x, y \in X$.

#BIS is conjectured to be located between **#SAT**-hard (no FPRAS unless RP=NP) and **FPRAS**able under **AP-reduction**.

Approximation-preserving reduction

M. E. Dyer, L. A. Goldberg, M. Jerrum, An approximation trichotomy for Boolean #CSP, J. Comput. Syst. Sci., 76(3-4): 267-277, 2010.



Intractable: What is #IDEALS?

Prob. #BIS

Given $G = (U, V; E)$ **B**ipartite graph.

Count the number of **I**ndependent **S**ets, where

$X \subseteq U \cup V$ is an *independent set* if $\{x, y\} \notin E$ for any $x, y \in X$.

Prob. #IDEALS

Given $\mathcal{P} = (N, \preccurlyeq)$ partially ordered set (poset).

Count the number of ideals, where

$X \subseteq N$ is an *ideal* if $x \in X$ and $y \preccurlyeq x$ then $y \in X$.

Thm.

#BIS has an FPRAS iff **#IDEALS** has an FPRAS.

Simply we say
"#IDEALS is #BIS-hard"

M. E. Dyer, L. A. Goldberg, C. S. Greenhill, M. Jerrum,
The relative complexity of approximate counting problems, Algorithmica 38(3), 471-500, 2004

Proof sketch: “#IDEALS is #BIS-hard”

SKIP this slide

If **#IDEALS** has an FPRAS then **#BIS** has an FPRAS.

Proof sketch.

□ If **#IDEALS** has an FPRAS then so does **#MaxBIS**.

Idea: Suppose $(A_1, B_1) \subseteq (U, V)$ and $(A_2, B_2) \subseteq (U, V)$ are respectively maximum independent sets of $G = (U, V; E)$. Then, both $(A_1 \cap A_2, B_1 \cup B_2)$ and $(A_1 \cup A_2, B_1 \cap B_2)$ are max. ind. set., meaning that it forms a distributive lattice w/appropriate meet/join. In fact, the representing poset is found in a polynomial time by Dulmage-Mendelsohn decomp.

□ If **#MaxBIS** has an FPRAS then so does **#BIS**.

Idea: By a Cook reduction (many-to-many).

Cf. M. E. Dyer, L. A. Goldberg, C. S. Greenhill, M. Jerrum, The relative complexity of approximate counting problems, Algorithmica 38(3): 471-500 (2004)

Proof sketch: "#BIS is #IDEALS-hard"

SKIP this slide

Conversely, if **#BIS** has an FPRAS then **#IDEALS** has an FPRAS.

Proof sketch.

☐ If **#BIS** has an FPRAS then so does **#MaxBIS**.

Idea: Let G' be a graph adding a pendant to every vertex in G. Then, ind. sets of G are bijective to max. ind. sets of G'.

☐ If **#MaxBIS** has an FPRAS then so does **#IDEALS**.

Idea: As given $\mathcal{P} = (N, \preccurlyeq)$, let $G = (U, V; E)$ be given by $|U| = |V| = |N|$ and $\{u_i, v_j\} \in E$ if $i \preccurlyeq j$. Then max. ind. sets of G are bijective to ideals of $\mathcal{P}$.

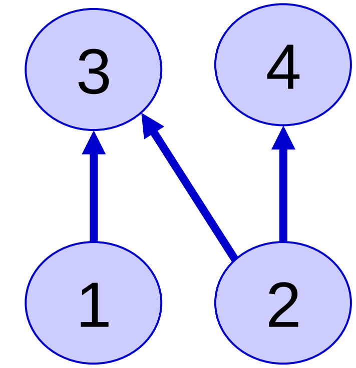
Cf. M. E. Dyer, L. A. Goldberg, C. S. Greenhill, M. Jerrum, The relative complexity of approximate counting problems, Algorithmica 38(3): 471-500 (2004)

Prob. #IDEALS

Given $\mathcal{P} = (N, \preccurlyeq)$ partially ordered set (poset).

Count the number of ideals, where

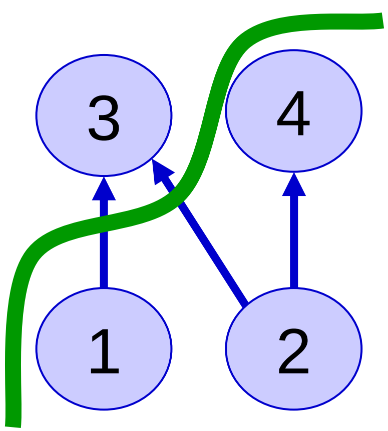
$X \subseteq N$ is an *ideal* if $x \in X$ and $y \preccurlyeq x$ then $y \in X$.



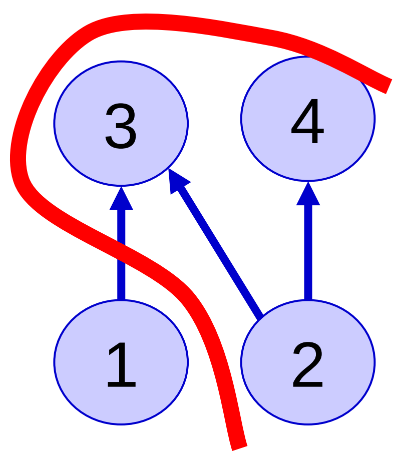
Let $\mathcal{J}(\mathcal{P}) = \{X \subseteq V \mid X \text{ is an ideal of } \mathcal{P}\}$.

$\mathcal{P} = (\{1,2,3,4\}, \preccurlyeq)$

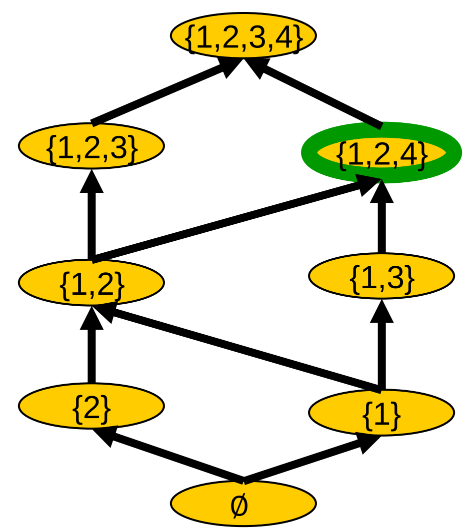
- ✓ $\mathcal{J}(\mathcal{P})$ forms a **distributive lattice** w.r.t. ∪ and ∩.
- ✓ Any finite **distributive lattice** is isomorphic to the set family of ideals of a poset (**Birkhoff's representation theorem**).



An ideal



Not an ideal



$\mathcal{J}(\mathcal{P})$

Intractable: log-supermodular is #BIS-hard

Prop. (representation by minimizers of a submodular fncs.)

As given a finite poset $\mathcal{P} = (N, \preccurlyeq)$, let $f: 2^N \to \mathbb{R}$ be given by

$$f(X) = |\{i \in X | \exists j \text{ such that } j \prec i \text{ and } j \notin X\}|$$

for $X \in 2^N$. Then f is submodular, and

$$f(X) \begin{cases} = 0 & \text{if } X \in \mathcal{J}(\mathcal{P}) \\ \geq 1 & \text{otherwise} \end{cases}$$

holds for $X \in 2^N$.

Let $g(X) = 2^{-(n+1)f(X)}$ for $X \in 2^N$, where $n = |N|$.

Notice that g is log-supermodular. Then

$$|\mathcal{J}(\mathcal{P})| \leq C \leq |\mathcal{J}(\mathcal{P})| + \frac{1}{2}$$

$$g(X) \begin{cases} = 1 & \text{if } X \in \mathcal{J}(\mathcal{P}) \\ \leq \frac{1}{2^{n+1}} & \text{otherwise} \end{cases}$$

where recall $C = \sum_{X \in 2^N} g(X)$. Thus $\lfloor C \rfloor = |\mathcal{J}(\mathcal{P})|$.

$\Rightarrow$ If we have an FPRAS for C we have an FPRAS for $|\mathcal{J}(\mathcal{P})|$.

Intermediate

- So far, we have seen that **sampling from log-supermodular distribution is #BIS-hard**, which is conjectured between #SAT (no FPRAS unless RP=NP) and FPRASable.

M. E. Dyer, L. A. Goldberg, M. Jerrum, An approximation trichotomy for Boolean #CSP, J. Comput. Syst. Sci., 76(3-4): 267-277, 2010.

- Why is it hard to sample from log-supermodular?

=> Two Hints(?)

1. Bad example for the simple Markov chain
2. log-supermodularity is not invariant under "transformation of variables".

Bad example for the simple Markov chain

Let $g: 2^N \to \mathbb{R}_{>0}$ be a log-supermodular fnc.

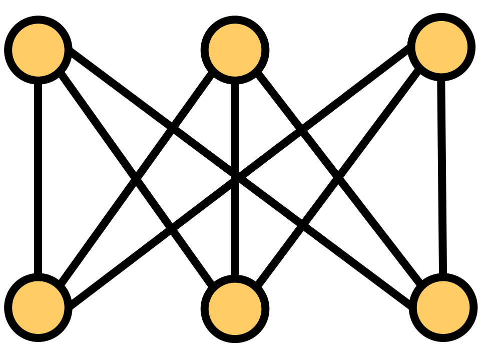
A transition from X to X' is defined as follows

1. Choose $i \in N$ u.a.r.
2. Let $X' = \begin{cases} X \cup \{i\} & \text{w.p.} \frac{g(X \cup \{i\})}{g(X \cup \{i\}) + g(X \setminus \{i\})}, \\ X \setminus \{i\} & \text{otherwise.} \end{cases}$

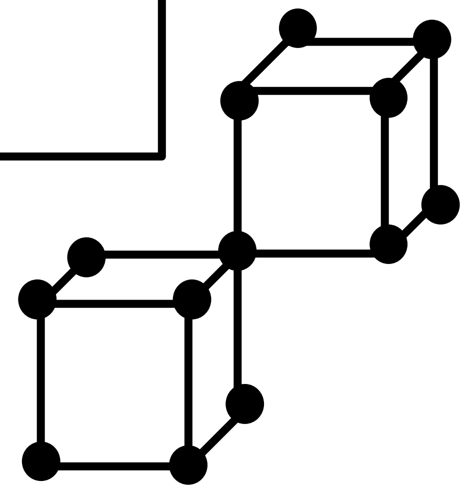
The mixing time of the MC $\geq 2^{\frac{n}{2}}$

The log-supermodular function for a poset $\mathcal{P} = (N, \preccurlyeq)$

$$g(X) \begin{cases} = 1 & \text{if } X \in \mathcal{J}(\mathcal{P}) \\ \leq \frac{1}{2^{n+1}} & \text{otherwise} \end{cases}$$



$\mathcal{P}$



$\mathcal{J}(\mathcal{P})$

Intermediate

- So far, we have seen that **sampling from log-supermodular distribution is #BIS-hard**, which is conjectured between #SAT (no FPRAS unless RP=NP) and FPRASable.

  M. E. Dyer, L. A. Goldberg, M. Jerrum, An approximation trichotomy for Boolean #CSP, J. Comput. Syst. Sci., 76(3-4): 267-277, 2010.

- Why is it hard to sample from log-supermodular?

=> Two Hints(?)

1. Bad example for the simple Markov chain
2. log-supermodularity is not invariant under "transformation of variables".

Goal of the talk

- ☐ Why log-supermodular?
  - ✓ Seemingly **"Tractable"**
    - ➢ Monotone coupling (cf. FKG ineq.), "log-concave" etc.
  - ✓ #BIS-**hard**
    - ➢ #Ideal, #stable matching

A challenge:
FPRAS or not

- ☐ What we (or I) know?
  - ✓ Log-concave?

    J. Nakashima, Y. Yamauchi, S. Kijima and M. Yamashita, Finding submodularity hidden in symmetric difference, SIAM Journal on Discrete Mathematics, 34:1 (2020), 571--585.
  - ✓ Subclass for #BIS-hard

    T. Fujii and S. Kijima, Every finite distributive lattice is isomorphic to the minimizer set of an $M^{\natural}$ -concave set function, Operations Research Letters, 49:1 (January 2021), 1--4.

SIAM J. DISCRETE MATH.
Vol. 34, No. 1, pp. 571–585

FINDING SUBMODULARITY HIDDEN IN SYMMETRIC DIFFERENCE*

JUNPEI NAKASHIMA†, YUKIKO YAMAUCHI†, SHUJI KIJIMA†, AND MASAFUMI YAMASHITA†

Abstract. A set function f on a finite set V is *submodular* if $f(X)+f(Y) \geq f(X \cup Y)+f(X \cap Y)$ for any pair $X, Y \subseteq V$. The *symmetric difference transformation* (*SD-transformation*) of f by a *canonical set* $S \subseteq V$ is a set function g given by $g(X) = f(X \triangle S)$ for $X \subseteq V$, where $X \triangle S = (X \setminus S) \cup (S \setminus X)$ denotes the symmetric difference between X and S. Submodularity and SD-transformations are regarded as the counterparts of convexity and affine transformations in a

submodular, although it requires exponentially many oracle calls in general.

Submodularity is often regarded as a discrete counter part of convexity,

- e.g., minimization is in P, maximization is NP-hard

Key words. submodular functions, symmetric difference

Convex functions

Def. Convex function

A function $f: \mathbb{R}^n \to \mathbb{R}$ is ***convex*** if

$$\lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \geq f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y})$$

holds for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and for any $\lambda \in [0,1]$.

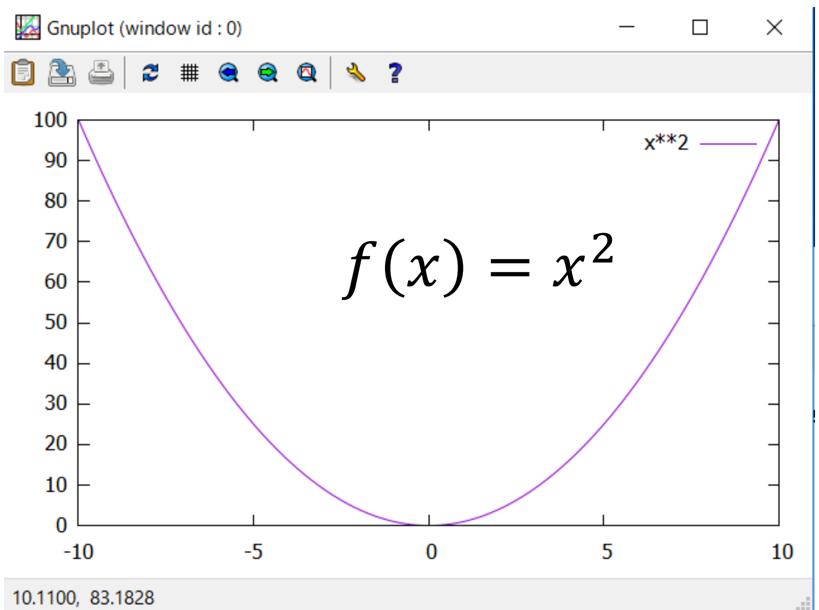
Convex functions

Def. Convex function

A function $f: \mathbb{R}^n \to \mathbb{R}$ is ***convex*** if

$$\lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \geq f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y})$$

holds for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and for any $\lambda \in [0,1]$.



convex

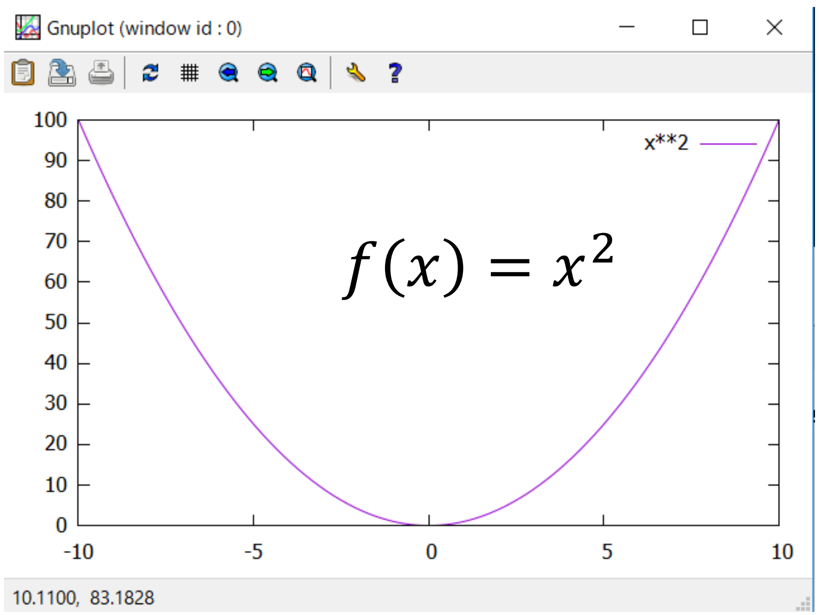
Convex functions

Def. Convex function

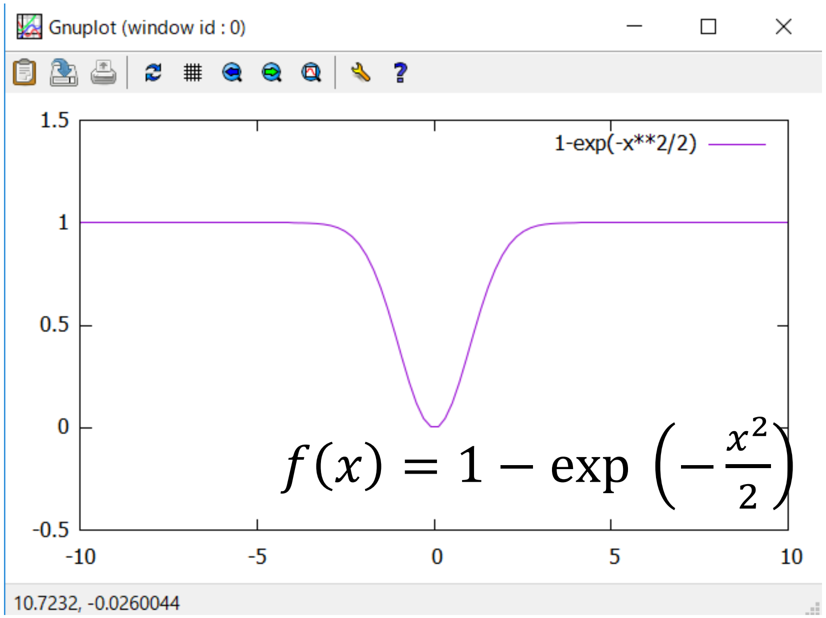
A function $f: \mathbb{R}^n \to \mathbb{R}$ is ***convex*** if

$$\lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \geq f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y})$$

holds for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and for any $\lambda \in [0,1]$.



convex



Non-convex

Exercises

Ex. 1.

$$f(x, y) = 2x^2 + 2xy + 5y^2$$

Is the function f convex?

Exercises

Ex. 1.

$$f(x,y) = 2x^2 + 2xy + 5y^2$$

Is the function f convex?

Answer.

Let

$$x \coloneqq s + 2t$$
$$y \coloneqq s - t.$$

Then,

$$\begin{aligned} g(s,t) &= f(s+2t, s-t) \\ &= 2(s+2t)^2 + 2(s+2t)(s-t) + 5(s-t)^2 \\ &= (2s^2 + 8st + 8t^2) + (2s^2 + 2st - 4t^2) + (5s^2 - 10st + 5t^2) \\ &= 9s^2 + 9t^2. \end{aligned}$$

Now, it is easy to observe that $g(s,t)$ is **convex**.

Exercises

Ex. 2.

$$f(x, y) = x^2 + 4xy + 3y^2$$

Is the function f convex?

Exercises

Ex. 2.

$$f(x, y) = x^2 + 4xy + 3y^2$$

Is the function f convex?

Answer.

Let

$$x := s + 3t$$
$$y := s - t.$$

Then,

$$\begin{aligned} g(s,t) &= f(s+3t, s-t) \\ &= (s+3t)^2 + 4(s+3t)(s-t) + 3(s-t)^2 \\ &= (s^2 + 6st + 9t^2) + (4s^2 + 8st - 12t^2) + (3s^2 - 6st + 3t^2) \\ &= 8s^2 + 8st \end{aligned}$$

$g(s,t)$ is **not convex**, that is confirmed by

e.g., $g(1, -2) = 8 - 16 = -8$, $g(-1,2) = 8 - 16 = -8$,

$g\left(\frac{1}{2}(1, -2) + \frac{1}{2}(-1,2)\right) = g(0,0) = 0 > \frac{1}{2}g(1, -2) + \frac{1}{2}g(-1,2).$

Convexity is invariant under affine transformation

Def. Convex function

A function $f: \mathbb{R}^n \to \mathbb{R}$ is ***convex*** if

$$\lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \geq f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y})$$

holds for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and for any $\lambda \in [0,1]$.

Thm. (cf. [Rockafellar])

Let $h: \mathbb{R}^n \to \mathbb{R}^n$ be an affine map.

If $f: \mathbb{R}^n \to \mathbb{R}$ is convex, then $g \coloneqq f \circ h$ is convex, too.

(i.e., let $h = A\boldsymbol{x} + \boldsymbol{b}$ where $A \in \mathbb{R}^{n \times n}$, $\boldsymbol{b} \in \mathbb{R}^n$,

then $g(\boldsymbol{x}) = f(A\boldsymbol{x} + \boldsymbol{b})$ is a convex function.

Convexity is invariant under affine transformation

Def. Convex function

A function $f: \mathbb{R}^n$ ...

holds f...

Main subject

Discrete analogy?

$\Rightarrow$ **submodular function**

Thm. (...

Let h: $\mathbb{R}$ ...

If $f: \mathbb{R}^n \to \mathbb{R}$ is co... ...vex, too.

(i.e., let $h = A\boldsymbol{x} + \boldsymbol{b}$ wh... $\in \mathbb{R}^{n \times n}$, $\boldsymbol{b} \in \mathbb{R}^n$,

then $g(\boldsymbol{x}) = f(A\boldsymbol{x} + \boldsymbol{b})$ is a convex function.

Def. submodular function

For a set function $f: \{0,1\}^V \to \mathbb{R}$, let $\Phi_f: \{0,1\}^V \times \{0,1\}^V \to \mathbb{R}$ be given by

$$\Phi_f(X,Y) := f(X) + f(Y) - f(X \cup Y) - f(X \cap Y).$$

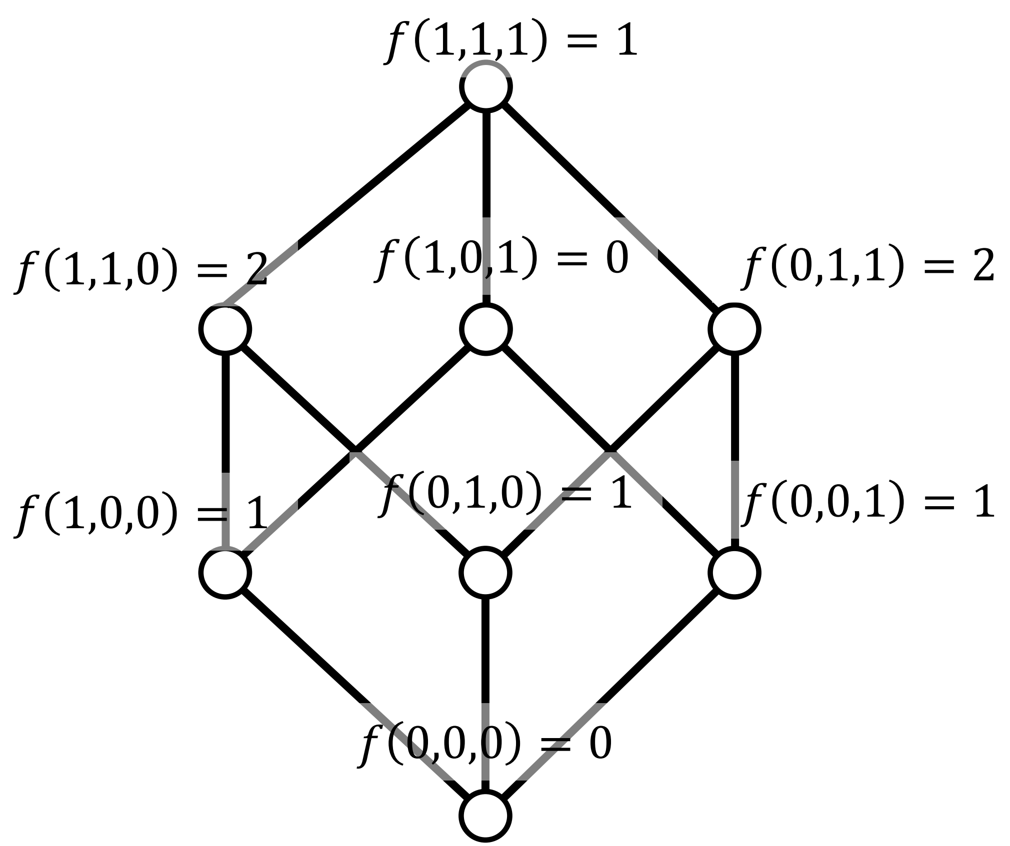
- A set function f is subdmodular if $\Phi_f(X,Y) \geq 0$ holds for any $X, Y \in \{0,1\}^V$.

What is natural for "*discrete variable transformation*"?

"Change origin" (+ rename)

Once assign **an origin**, a Boolean lattice is *uniquely* determined (except for the *name* of items).

We describe an "assignment of an origin" by a ***symmetric difference transformation***, in the next slides.



Def. submodular function

For a set function $f\colon \{0,1\}^V \to \mathbb{R}$, let $\Phi_f\colon \{0,1\}^V \times \{0,1\}^V \to \mathbb{R}$ be given by

$$\Phi_f(X,Y) \coloneqq f(X) + f(Y) - f(X \cup Y) - f(X \cap Y).$$

- A set function f is subdmodular if $\Phi_f(X,Y) \geq 0$ holds for any $X, Y \in \{0,1\}^V$.

symmetric difference transformation

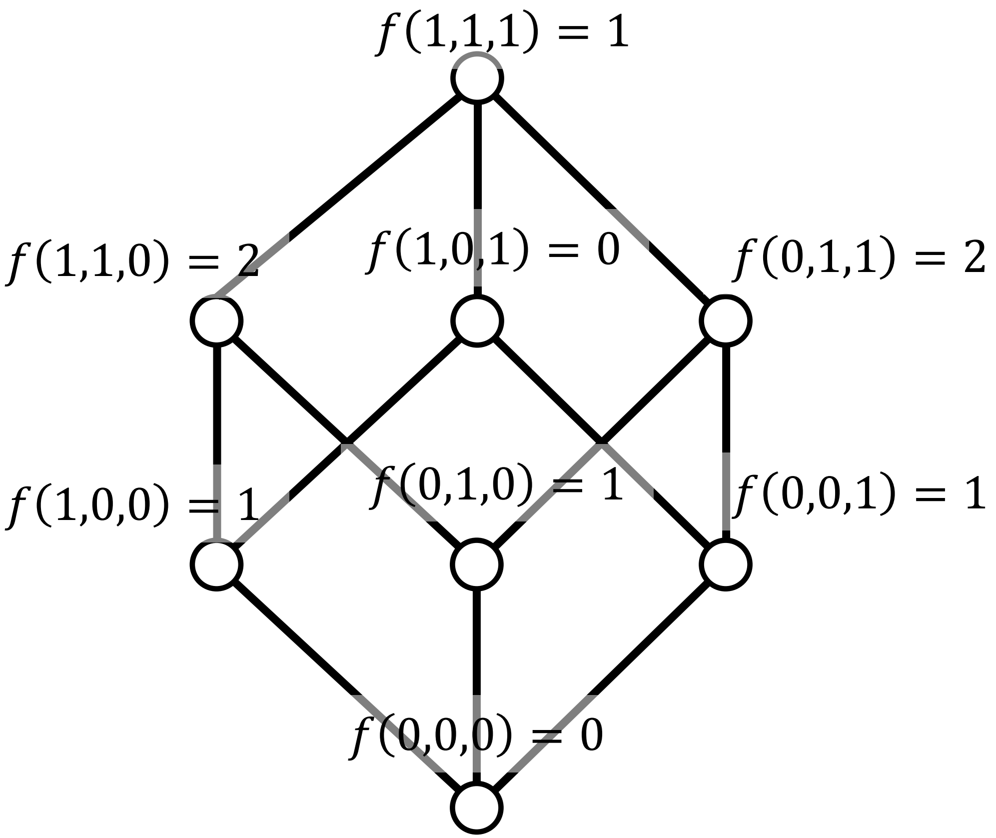
Def. SD-transformation

Let $\sigma_S\colon \{0,1\}^V \to \{0,1\}^V$ ($S \subseteq V$) be a SD-map given by

$$\sigma_S(X) \coloneqq X \oplus S \quad (X \in \{0,1\}^V).$$

The "*SD-transformation* of f by S" is a set function $g\colon \{0,1\}^V \to \mathbb{R}$ defined by

$$g \coloneqq f \circ \sigma_S.$$



Def. submodular function

For a set function $f: \{0,1\}^V \to \mathbb{R}$, let $\Phi_f: \{0,1\}^V \times \{0,1\}^V \to \mathbb{R}$ be given by

$$\Phi_f(X,Y) := f(X) + f(Y) - f(X \cup Y) - f(X \cap Y).$$

- A set function f is subdmodular if $\Phi_f(X,Y) \geq 0$ holds for any $X, Y \in \{0,1\}^V$.

symmetric difference transformation

Def. SD-transformation

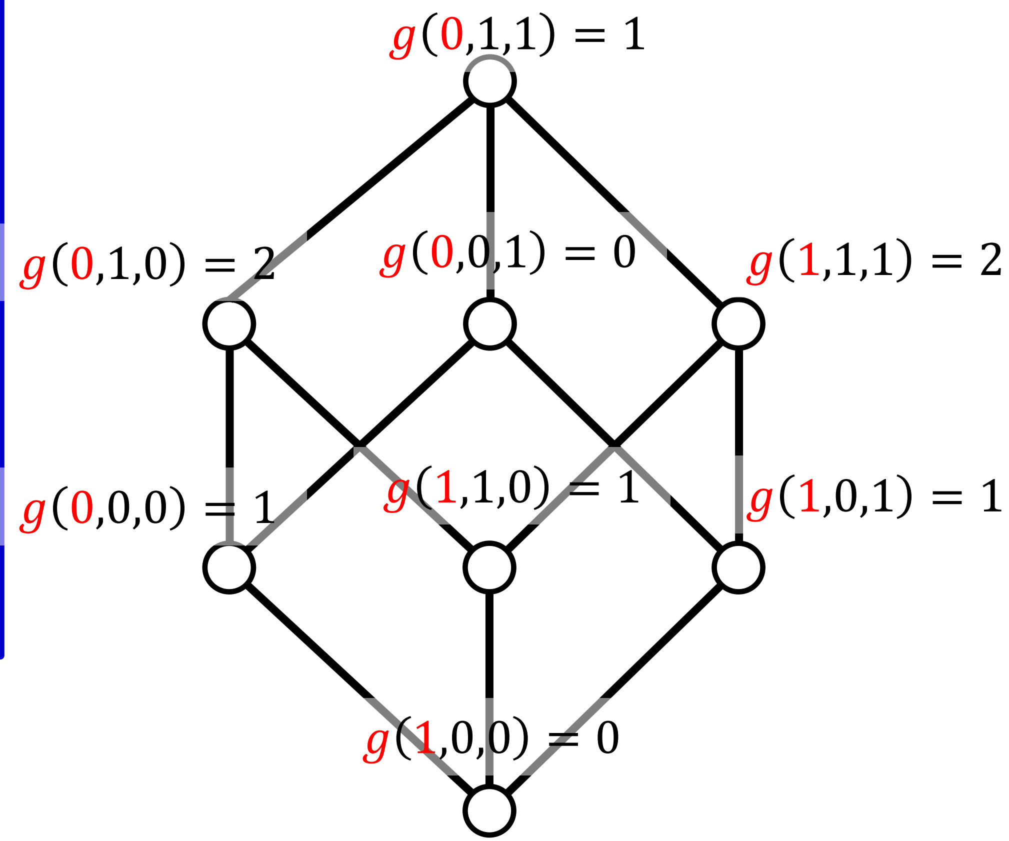
Let $\sigma_S: \{0,1\}^V \to \{0,1\}^V$ ($S \subseteq V$) be a SD-map given by

$$\sigma_S(X) := X \oplus S \quad (X \in \{0,1\}^V).$$

The "*SD-transformation* of f by S" is a set function $g: \{0,1\}^V \to \mathbb{R}$ defined by

$$g := f \circ \sigma_S.$$

Ex. $g := f \circ \sigma_{(1,0,0)}$



Def. submodular function

For a set function $f: \{0,1\}^V \to \mathbb{R}$, let $\Phi_f: \{0,1\}^V \times \{0,1\}^V \to \mathbb{R}$ be given by

$$\Phi_f(X,Y) := f(X) + f(Y) - f(X \cup Y) - f(X \cap Y).$$

- A set function f is subdmodular if $\Phi_f(X,Y) \geq 0$ holds for any $X, Y \in \{0,1\}^V$.

symmetric difference transformation

Def. SD-transformation

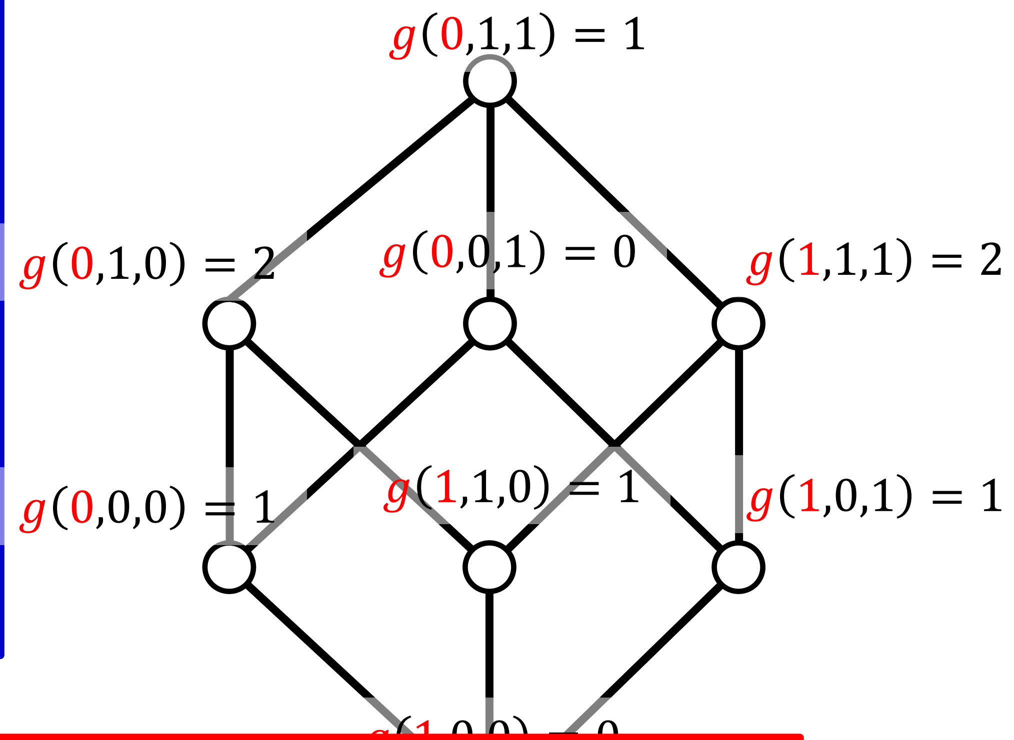
Let $\sigma_S: \{0,1\}^V \to \{0,1\}^V$ ($S \subseteq V$) be a SD-map given by

$$\sigma_S(X) := X \oplus S \quad (X \in \{0,1\}^V).$$

The "*SD-transformation* of f by S" is a set function $g: \{0,1\}^V \to \mathbb{R}$ defined by

$$g := f \circ \sigma_S.$$

Ex. $g := f \circ \sigma_{(1,0,0)}$



SD-transformations do not inherit submodularity, in general

<u>Def. submodular function</u>

For a set function $f:\{0,1\}^V \to \mathbb{R}$, let $\Phi_f:\{0,1\}^V \times \{0,1\}^V \to \mathbb{R}$ be given by

$$\Phi_f(X,Y) := f(X) + f(Y) - f(X \cup Y) - f(X \cap Y).$$

- A set function f is subdmodular if $\Phi_f(X,Y) \geq 0$ holds for any $X, Y \in \{0,1\}^V$.

symmetric difference transformation

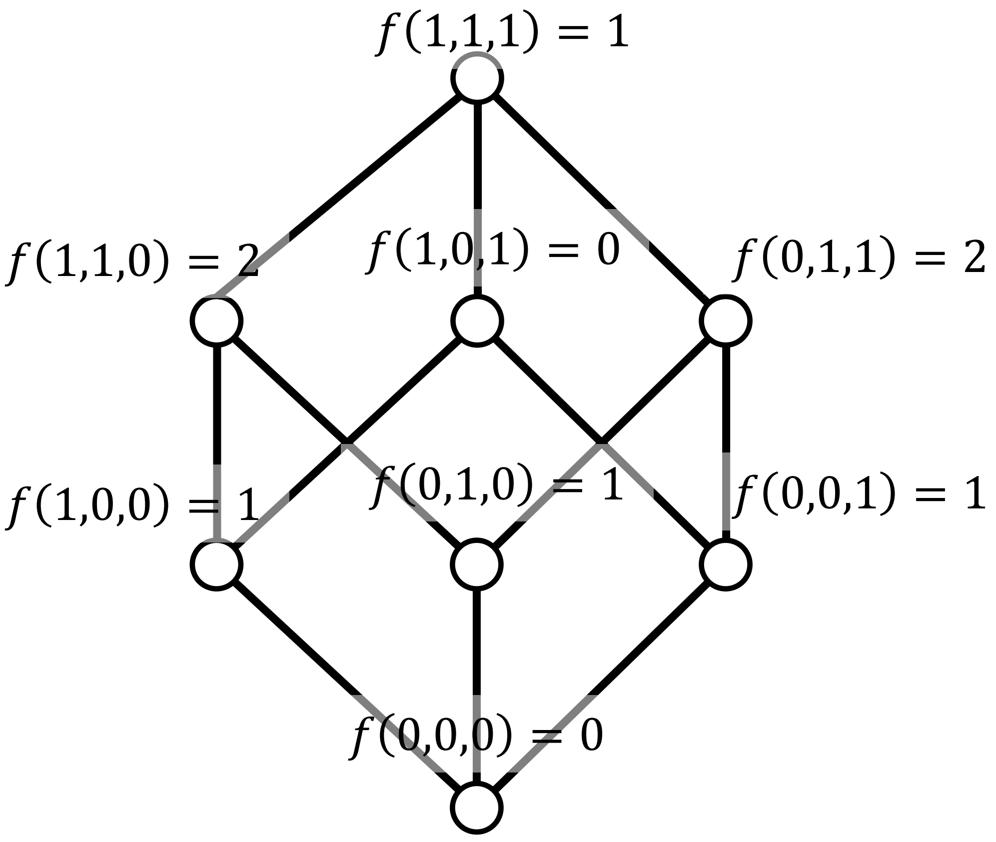
<u>Def. SD-transformation</u>

Let $\sigma_S:\{0,1\}^V \to \{0,1\}^V$ ($S \subseteq V$) be a SD-map given by

$$\sigma_S(X) := X \oplus S \quad (X \in \{0,1\}^V).$$

The "*SD-transformation* of f by S" is a set function $g:\{0,1\}^V \to \mathbb{R}$ defined by

$$g := f \circ \sigma_S.$$



Def. submodular function

For a set function $f: \{0,1\}^V \to \mathbb{R}$, let $\Phi_f: \{0,1\}^V \times \{0,1\}^V \to \mathbb{R}$ be given by

$$\Phi_f(X,Y) \coloneqq f(X) + f(Y) - f(X \cup Y) - f(X \cap Y).$$

- A set function f is subdmodular if $\Phi_f(X,Y) \geq 0$ holds for any $X, Y \in \{0,1\}^V$.

symmetric difference transformation

Def. SD-transformation

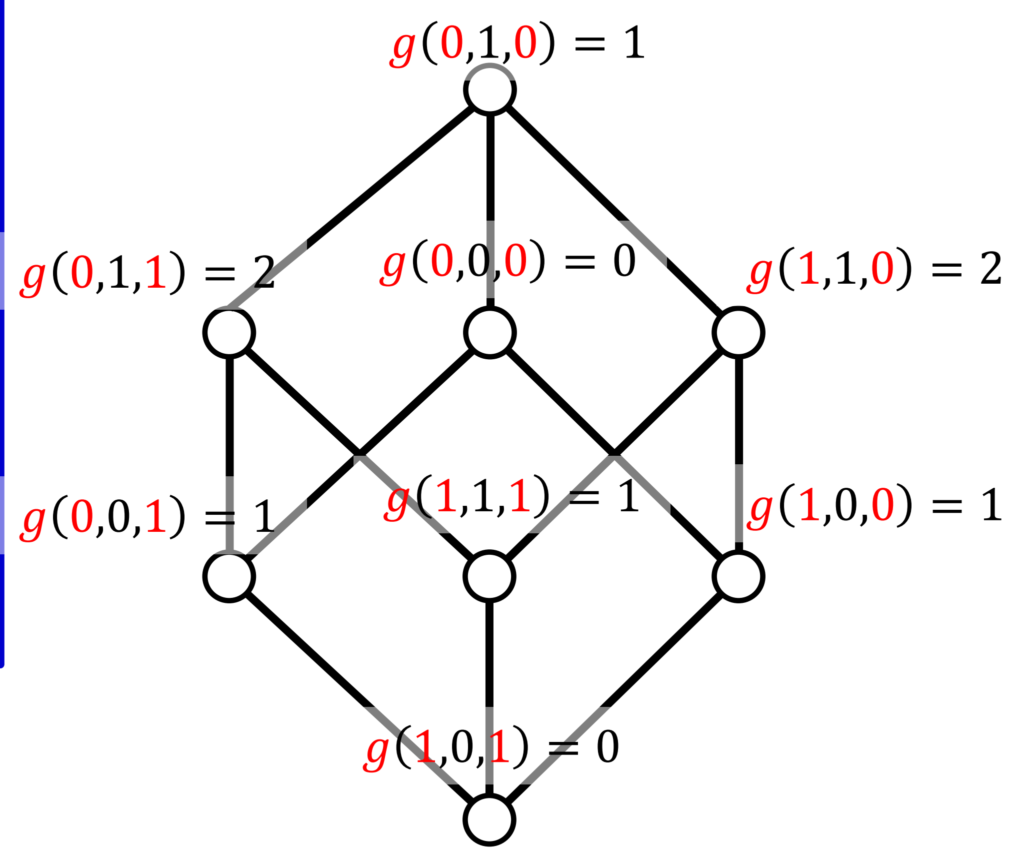
Let $\sigma_S: \{0,1\}^V \to \{0,1\}^V$ ($S \subseteq V$) be a SD-map given by

$$\sigma_S(X) \coloneqq X \oplus S \quad (X \in \{0,1\}^V).$$

The "*SD-transformation* of f by S" is a set function $g: \{0,1\}^V \to \mathbb{R}$ defined by

$$g \coloneqq f \circ \sigma_S.$$

Ex. $g \coloneqq f \circ \sigma_{(1,0,1)}$



Def. submodular function

For a set function $f: \{0,1\}^V \to \mathbb{R}$, let $\Phi_f: \{0,1\}^V \times \{0,1\}^V \to \mathbb{R}$ be given by

$$\Phi_f(X,Y) := f(X) + f(Y) - f(X \cup Y) - f(X \cap Y).$$

- A set function f is subdmodular if $\Phi_f(X,Y) \geq 0$ holds for any $X, Y \in \{0,1\}^V$.

symmetric difference transformation

Def. SD-transformation

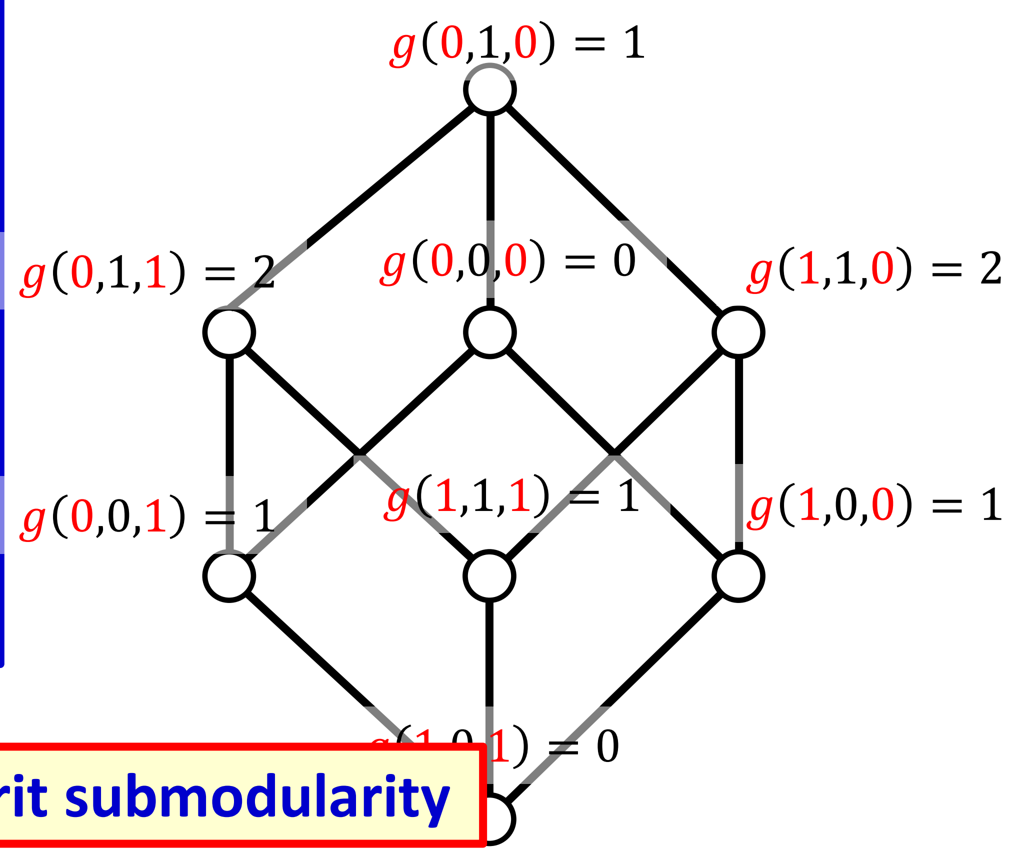
Let $\sigma_S: \{0,1\}^V \to \{0,1\}^V$ ($S \subseteq V$) be a SD-map given by

$$\sigma_S(X) := X \oplus S \quad (X \in \{0,1\}^V).$$

The "*SD-transformation* of f by S" is a set function $g: \{0,1\}^V \to \mathbb{R}$ defined by

$$g := f \circ \sigma_S.$$

Ex. $g := f \circ \sigma_{(1,0,1)}$



Some SD-transformations inherit submodularity

Def. submodular function

For a set function $f: \{0,1\}^V \to \mathbb{R}$, let $\Phi_f: \{0,1\}^V \times \{0,1\}^V \to \mathbb{R}$ be given by

$$\Phi_f(X,Y) := f(X) + f(Y) - f(X \cup Y) - f(X \cap Y).$$

- A set function f is subdmodular if $\Phi_f(X,Y) \geq 0$ holds for any $X, Y \in \{0,1\}^V$.

symmetric difference transformation

When does S preserve submodularity?

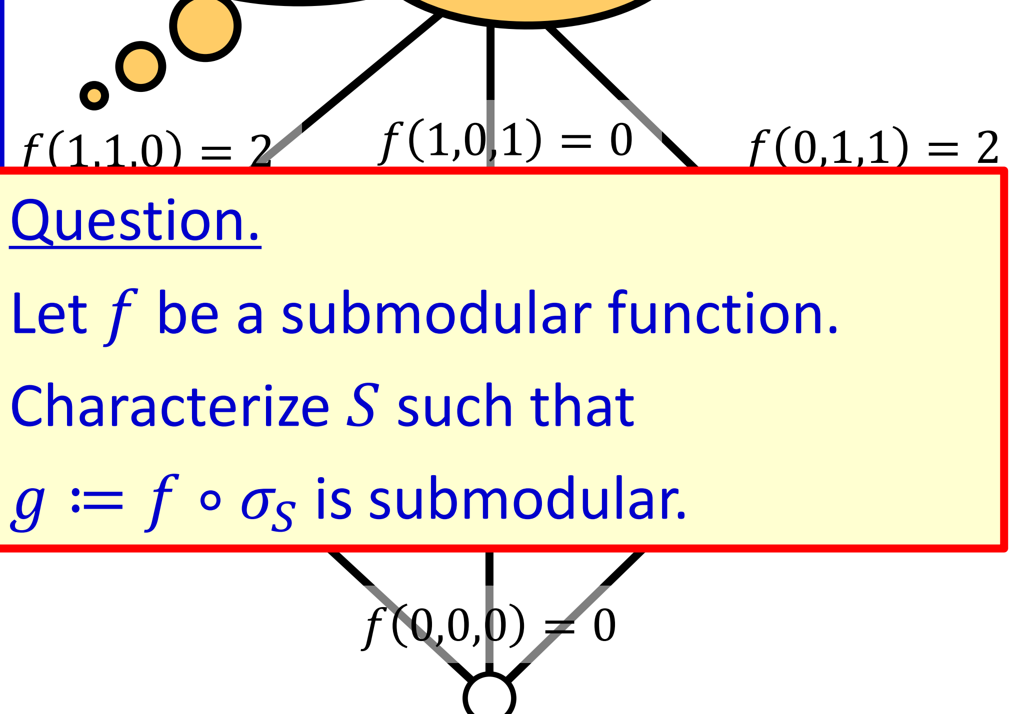
Def. SD-transformation

Let $\sigma_S: \{0,1\}^V \to \{0,1\}^V$ ($S \subseteq V$) be a SD-map given by

$$\sigma_S(X) := X \oplus S \quad (X \in \{0,1\}^V).$$

The "*SD-transformation* of f by S" is a set function $g: \{0,1\}^V \to \mathbb{R}$ defined by

$$g := f \circ \sigma_S.$$



Question.

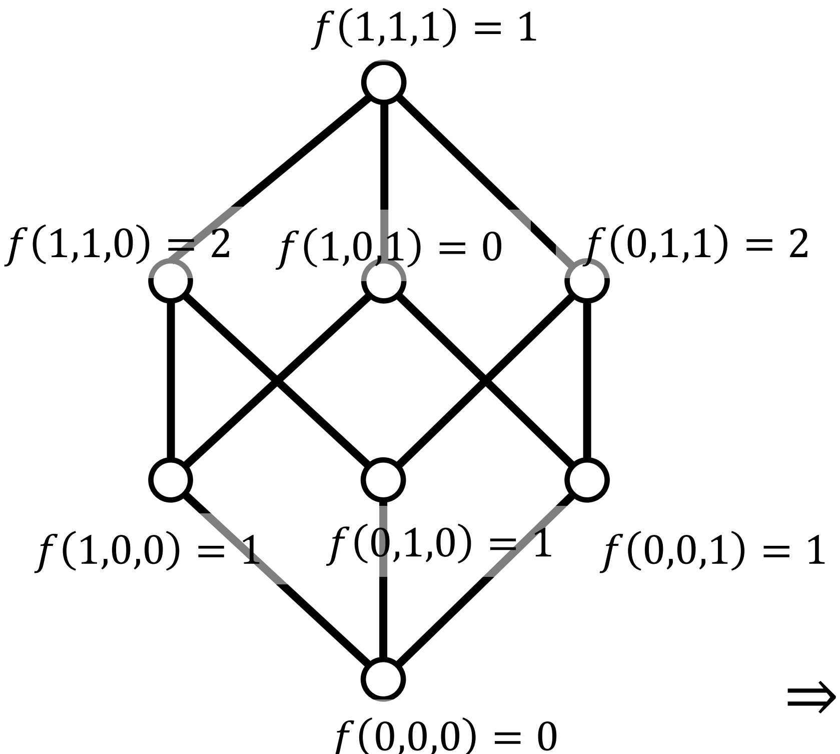
Let f be a submodular function. Characterize S such that $g := f \circ \sigma_S$ is submodular.

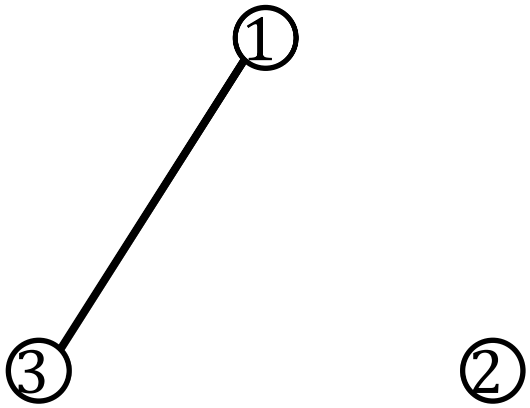
Main Result

Thm. 2 (Main Thm.)

For any $S \subseteq V$,
$[f \circ \sigma_S \text{ is submodular} \iff S \in \mathcal{U}(f)]$

Let
$\mathcal{U}(f) := \{ \bigcup_{i \in I} U_i \mid I \subseteq \{1, \dots, k\} \}$
where $U_i \subseteq V$ $(i = 1, \dots, k)$ is con. comp.
(k is #con. comp. of G_f)





$\Rightarrow \mathcal{U}(f) = \{\emptyset, \{2\}, \{1,3\}, \{1,2,3\}\}.$

Intermediate

- So far, we have seen that **sampling from log-supermodular distribution is #BIS-hard**, which is conjectured between #SAT (no FPRAS unless RP=NP) and FPRASable.

M. E. Dyer, L. A. Goldberg, M. Jerrum, An approximation trichotomy for Boolean #CSP, J. Comput. Syst. Sci., 76(3-4): 267-277, 2010.

- Why is it hard to sample from log-supermodular?

=> Two Hints(?)

1. Bad example for the simple Markov chain
2. log-supermodularity is not invariant under "transformation of variables".

Almost end

A confe e...

- Kazuo Murota said ...

“Kijima, do you know *$M^{\#}$-concave set functions* form a proper subclass of submodular fns.”

(So, sampling from log-$M^{\#}$-convex distributions may be easier than from log-supermodular distr., as I understand)

- Both Min/Maximization of an $M^{\#}$-concave fn. is in P.

(It looks like a matroid rank function, but not “monotone”)

Answer: Sampling from log-$M^{\#}$-convex distribution is still hard.

T. Fujii and S. Kijima, Any finite distributive lattice is isomorphic to the minimizer set of an $M^{\natural}$-concave set function, arXiv 1903.08343, 2019.

Goal of the talk

- Why log-supermodular?
  - ✓ Seemingly **“Tractable”**
    - ➢ Monotone coupling (cf. FKG ineq.), “log-concave” etc.
  - ✓ #BIS-**hard**
    - ➢ #Ideal, #stable matching

A challenge:
FPRAS or not

- What we (or I) know?
  - ✓ Log-concave?

    J. Nakashima, Y. Yamauchi, S. Kijima and M. Yamashita, Finding submodularity hidden in symmetric difference, SIAM Journal on Discrete Mathematics, 34:1 (2020), 571--585.

  - ✓ Subclass for #BIS-hard

    T. Fujii and S. Kijima, Every finite distributive lattice is isomorphic to the minimizer set of an $M^{\natural}$-concave set function, Operations Research Letters, 49:1 (January 2021), 1--4.

Operations Research Letters 49 (2021) 1–4

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Every finite distributive lattice is isomorphic to the minimizer set of an M$^\natural$-concave set function

Tomohito Fujii, Shuji Kijima *

Graduate School of Information Science and Electronic Engineering, Kyushu University, Japan

ARTICLE INFO

Article history:
Received 4 May 2020
Received in revised form 31 August 2020
Accepted 28 October 2020
Available online 3 November 2020

Keywords:
M$^\natural$-concave
Valuated matroid
Submodular
Distributive lattice

ABSTRACT

$M^\natural$-concavity is a key concept in discrete convex analysis. For set functions, the class of $M^\natural$-concavity is a proper subclass of submodularity. It is a well-known fact that the set of minimizers of a submodular function forms a distributive lattice, where every finite distributive lattice is possible to appear. It is a natural question whether every finite distributive lattice appears as the minimizer set of an $M^\natural$-concave set function. This paper affirmatively answers the question.

1. Introduction

Proposition 2 (*See e.g.,* [9]). *Given a finite poset* $\mathcal{P} = (N, \preccurlyeq)$, *let*

Formal Definition of f (1/2)

Let $\mathcal{P} = (N, \preccurlyeq)$ be a poset and $N \coloneqq \{1,2,\dots,n\}$.

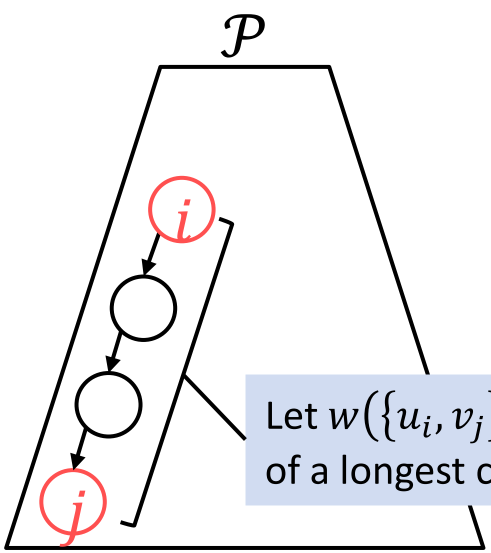
We define a weighted bipartite graph $G \coloneqq (U, V; E)$ as follows.

The vertex set is given by the union of $U \coloneqq \{u_1, u_2, \dots, u_n\}$ and $V \coloneqq \{v_1, v_2, \dots, v_n\}$.

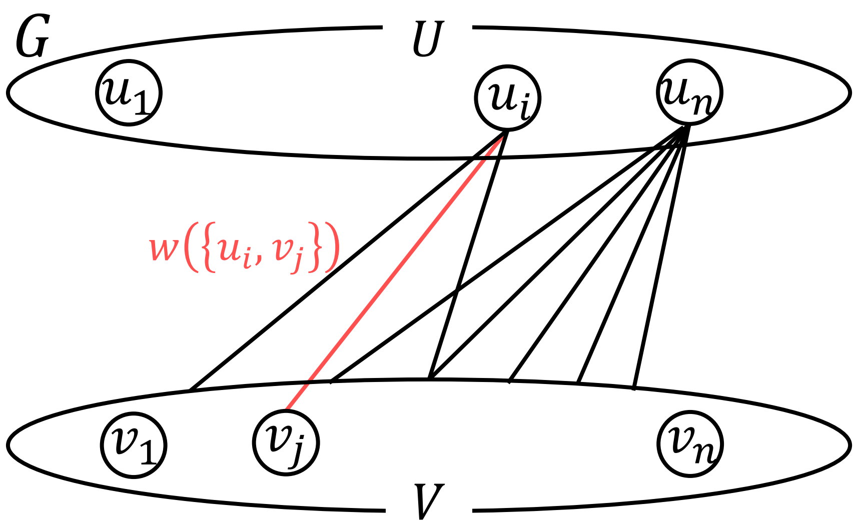
The edge set $E \coloneqq \left\{\{u_i, v_j\} : u_i \in U, v_j \in V \text{ and } j \prec i \text{ on } \mathcal{P}\right\}$.

The edge weight $w : E \to \mathbb{Z}_{\geq 0}$ is given by

$$w(\{u_i, v_j\}) = \max\{|X| - 1 : X \subseteq N \text{ is a chain between } j \text{ and } i\}.$$



Let $w(\{u_i, v_j\})$ be the length of a longest chain from i to j.



Formal Definition of f (2/2)

We define a set function $f: 2^N \to \mathbb{R}$ by

$$f(X) = \max\left\{\sum_{e \in M} w(e) : M \text{ is a matching between } U_X \text{ and } V_{\bar{X}}\right\}$$

where $U_X \coloneqq \{u_i \in U : i \in X\}$ and $V_{\bar{X}} \coloneqq \{v_i \in V : i \notin X\}$.

Thm.

f is M$^\#$-concave, and its minimizer set is $\mathcal{I}(\mathcal{P})$.

Cor.

If we have an efficient sampler from log-M$^\#$-convex distribution, then we have an FPRAS for #IDEALS, and hence for #BIS.

Goal of the talk

- Why log-supermodular?
  - ✓ Seemingly **"Tractable"**
    - ➢ Monotone coupling (cf. FKG ineq.), "log-concave" etc.
  - ✓ #BIS-**hard**
    - ➢ #Ideal, #stable matching

A challenge:
FPRAS or not

- What we (or I) know?
  - ✓ Log-concave?

    J. Nakashima, Y. Yamauchi, S. Kijima and M. Yamashita, Finding submodularity hidden in symmetric difference, SIAM Journal on Discrete Mathematics, 34:1 (2020), 571--585.

  - ✓ Subclass for #BIS-hard

    T. Fujii and S. Kijima, Every finite distributive lattice is isomorphic to the minimizer set of an $M^{\natural}$-concave set function, Operations Research Letters, 49:1 (January 2021), 1--4.

Conclusion

For efficient sampling from an arbitrary log-supermodular distribution, #BIS needs to have an FPRAS.

Some people conjecture that #BIS is located between #SAT (no FPRAS, in general) and FPRASable.

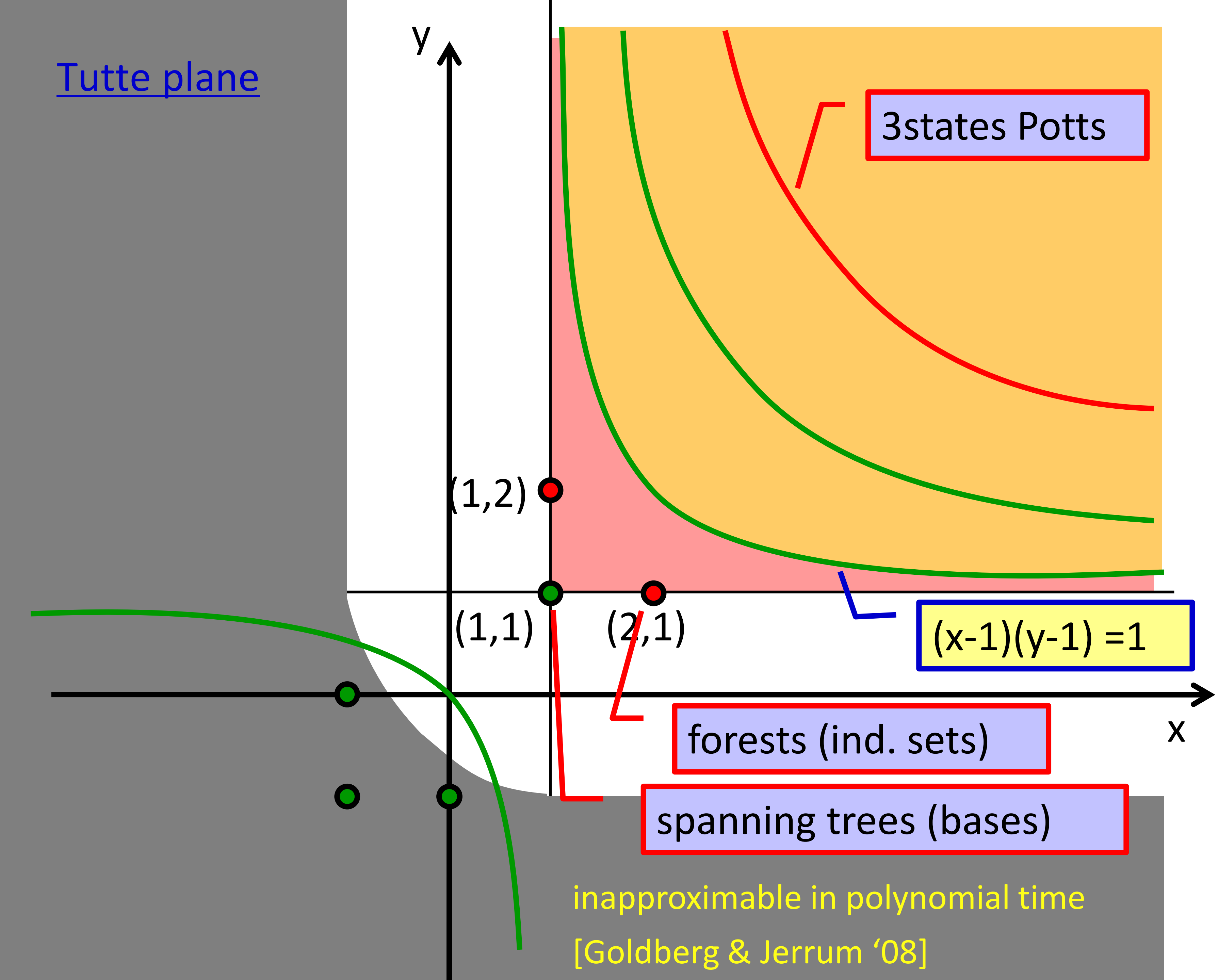
- ✓ Bad News: sampling from log-$M^{\#}$-concave distribution is still #BIS-hard.
- ✓ Good News: #BIS (#IDEALS, precisely) is restricted to log-$M^{\#}$-concave distribution, from log-supermodular distribution.

Good News? Recent development

N. Anari, K. Liu, S. O. Gharan, C. Vinzant, Log-concave polynomials II: high-dimensional walks and an FPRAS for counting bases of a matroid, STOC '19.

Provides an FPRAS for $T_G(2,1)$ ($T_M(1,1)$ and $T_M(2,1)$, in fact)

Tutte plane


y
x
(1,2)
(1,1)
(2,1)
3states Potts
(x-1)(y-1) =1
forests (ind. sets)
spanning trees (bases)
inapproximable in polynomial time
[Goldberg & Jerrum '08]

The end

Thank you for the attention.