Computing the nc-rank via discrete convex optimization on CAT(0) spaces

Hiroshi Hirai
The University of Tokyo
JST PRESTO

Joint work with Masaki Hamada

SIAM Conference on Applied Algebraic Geometry (AG21)
August 16, 2021

Edmonds Problem Edmonds 1967

Can we compute the rank of

$$A = A_1 x_1 + A_2 x_2 + \cdots + A_m x_m$$

in polynomial time ?

x_i : variables, A_i: $n \times n$ matrices over field $\mathbb{K}$

A: matrix over $\mathbb{K}[x_1, x_2, \ldots, x_m] \subset \mathbb{K}(x_1, x_2, \ldots, x_m)$

- RP, but P ? (for large field)
- Related to fundamental problems in diverse areas
  ～ combinatorial optimization, rigidity theory, TCS,…

Non-commutative Edmonds Problem

Ivanyos-Qiao-Subrahmanyam 2017

Can we compute the rank ***(nc-rank)*** of

$$A = A_1 x_1 + A_2 x_2 + \cdots + A_m x_m$$

in polynomial time ?

x_i : noncommutative variables, A_i: matrices over field $\mathbb{K}$

A: matrix over free ring $\mathbb{K}\langle x_1, x_2, \ldots, x_m \rangle$

$\cap$

free skew field $\mathbb{K}(\langle x_1, x_2, \ldots, x_m \rangle)$

- Broaden the literature:
  Noncommutative algebra, Invariant theory, …

Null cone membership for left-right action
$GL_n \times GL_n \ni (S, T) \mapsto (SA_k T)_{k=1,2,\ldots,m}$

nc-rank in **P**

- Garg-Gurvits-Oliveira-Wigderson 2019 : $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$

  Gurvits' operator scaling $\approx$ alternating minimization for a *geodesically convex* function on symmetric space GL_n/O_n

  nc-nonsingularity $\Leftrightarrow$ $\inf \left\{ \log \frac{\det \sum A_i X A_i^T}{\det X} : X \succ 0 \right\} > -\infty$

  → *Noncommutative optimization* (Bürgisser et al.)

- Ivanyos-Qiao-Subrahmanyam 2018: $\mathbb{K}$ arbitrary

  Wong sequence --- vector-space analogue of augmenting path

  Matrix completion by cyclic division algebra

Our contribution [Hamada-H 2020]

A different polytime algorithm to compute nc-rank

Features / techniques:

- Submodular optimization on modular lattice

$$f(p) + f(q) \geq f(p \wedge q) + f(p \vee q)$$

- Geodesically convex optimization on CAT(0) space

non-manifold

For nc-rank over $\mathbb{Q}$ (to avoid bit-complexity issue):

- reduction to nc-rank over $\mathrm{GF}(p)$ by p-adic valuation

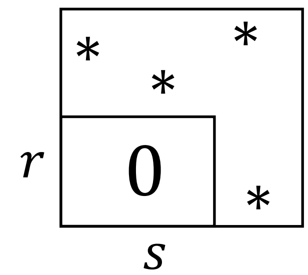
→ discrete convex optimization

on Euclidean building for $GL_n(\mathbb{Q})$

Submodularity in nc-rank

Thm (Fortin-Reutenauer 2004)

$$\text{nc-rank} \sum_k A_k x_k = 2n - \text{Max. } r + s$$

s.t. $PA_kQ =$ 

$(\forall k)$

$P, Q \in \mathrm{GL}_n(\mathbb{K})$

$$\text{Max. } \dim X + \dim Y \quad \text{s.t. } A_k(X, Y) = \{0\} \ (\forall k)$$

$X, Y \subseteq \mathbb{K}^n$ vector subspaces

where $A_k(x, y) \coloneqq x^T A_k y$

Max. $\dim X + \dim Y$ s.t. $A_k(X,Y) = 0\ (\forall k)$

$X, Y \subseteq \mathbb{K}^n$ vector subspaces

$2n+1$

$\cong$

Min. $-\dim X - \dim Y + C \sum_k \operatorname{rank} A_k|_{X \times Y}$

s.t. $(X,Y) \in \mathcal{L} \times \mathcal{L}$

$\mathcal{L}$: the modular lattice of all vector subspaces of $\mathbb{K}^n$

partial order reversed

Lem: The objective function is submodular in $\mathcal{L} \times \mathcal{L}^*$

$$f(X,Y) + f(X',Y') \geq f(X+X', Y \cap Y') + f(X \cap X', Y + Y')$$

$\wedge = (+, \cap),\ \vee = (\cap, +)$

Our approach to solve SFM for nc-rank

$$\text{Min.}\quad -\dim X - \dim Y + C \textstyle\sum_i \operatorname{rank} A_i|_{X\times Y}$$

$$\text{s.t.}\ (X, Y) \in \mathcal{M} = \mathcal{L} \times \mathcal{L}^*$$

- Continuous & convex relaxation to *CAT(0) space*

$\mathcal{M} \overset{\text{orthoscheme complex}}{\Longrightarrow} K(\mathcal{M})$, a CAT(0) space

$f \overset{\text{Lovász extension}}{\Longrightarrow} \bar{f}: K(\mathcal{M}) \to \mathbb{R}$, geodesically convex

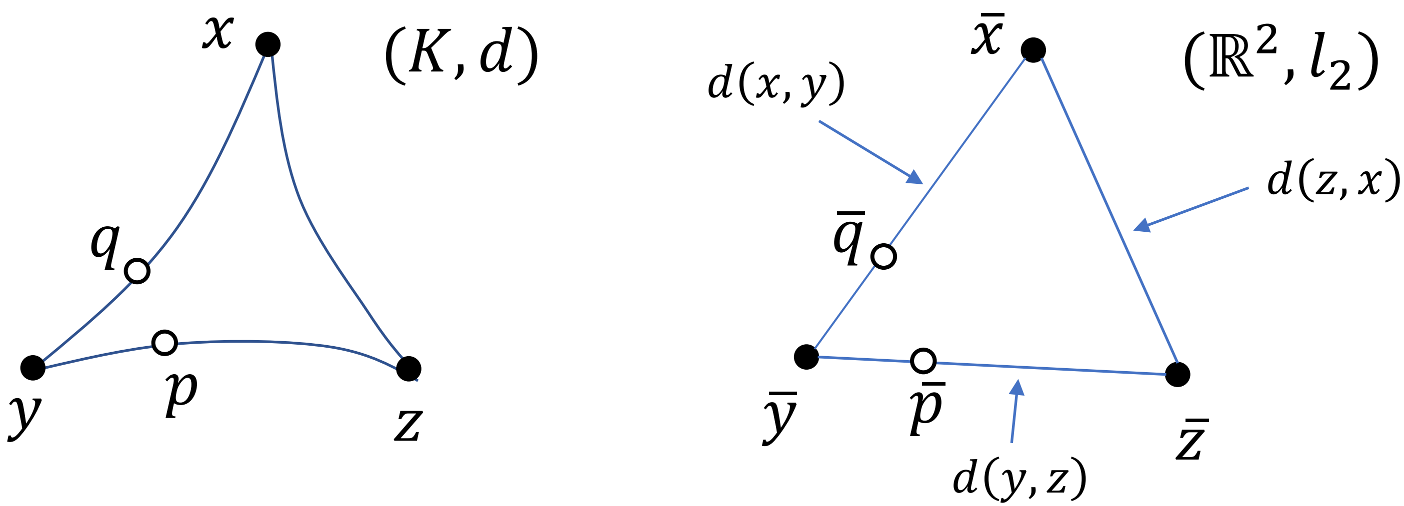
- Minimize $\bar{f}$ by an optimization algorithm on CAT(0) space

Inspired by the classical SFM [Lovász-Grötschel-Schrijver] :

- $f: \{0,1\}^n \to \mathbb{R}$, submodular $\Rightarrow$ $\bar{f}: [0,1]^n \to \mathbb{R}$, convex
- Minimize $\bar{f}$ by ellipsoid method

CAT(0) space (Gromov 1987)

A geodesic metric space (K, d) s.t. every triangle is "slimmer"



CAT(0)-inequality: $d(p, q) \leq \|\bar{p} - \bar{q}\|_2$

Ex: $\mathbb{R}^n$, hyperbolic space, symmetric space of noncompact type

tree, Euclidean building, …

Modeling & Algorithm & Optimization on CAT(0) spaces

Fact: CAT(0) space is uniquely geodesic

→ Convexity is naturally defined

- The space of phylogenetic trees (BHV-tree space) [Billera et al 2001]
- Configuration spaces of robots [Abram &Ghrist 2004]
- Polynomial-time geodesic computation

  for BHV-tree space [Owen 2011]

  for CAT(0) cubical complex [Hayashi 2021]
- Proximal point algorithm for convex optimization [Bačák 2014]
- Combinatorially-defined CAT(0) spaces:

  median graph → CAT(0) cubical complex [Chepoi 2000]

  poset / lattice → orthoscheme complex [Brady-McCammond 2012]

Orthoscheme Complex Brady-McCammond 2012

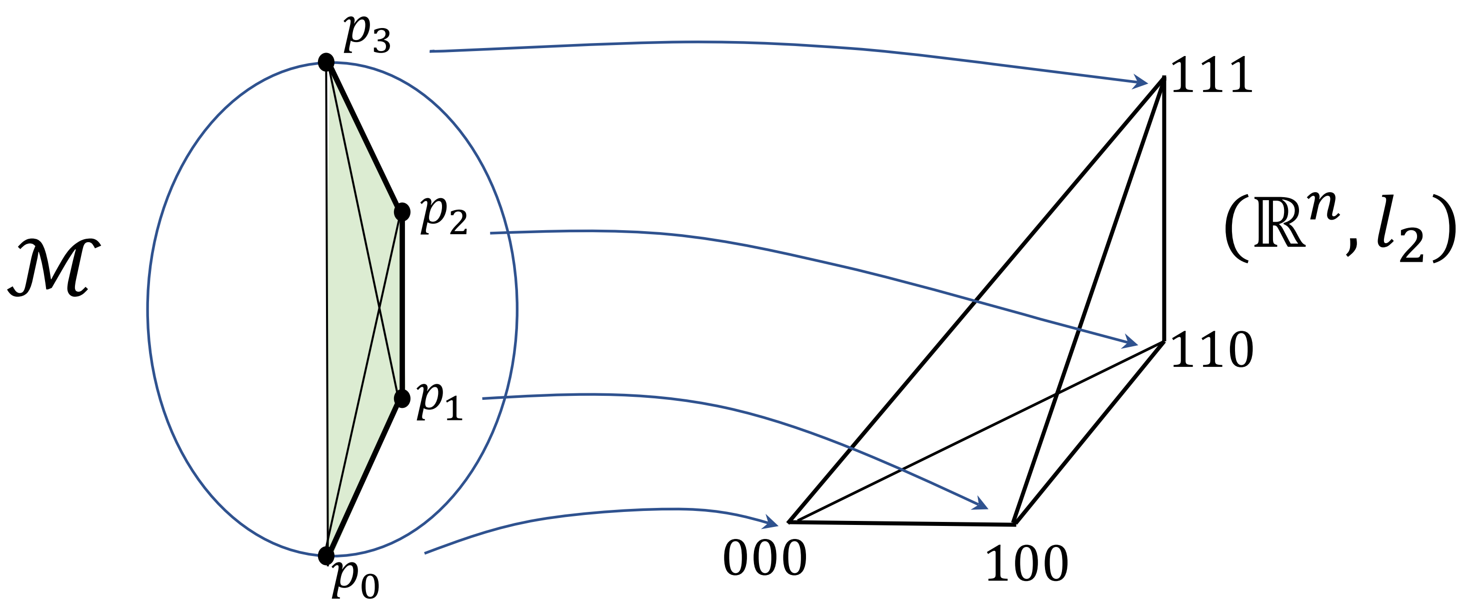
$\mathcal{M}$: graded poset

formal convex combination $\sum_i \mu_i = 1, \mu_i \geq 0$

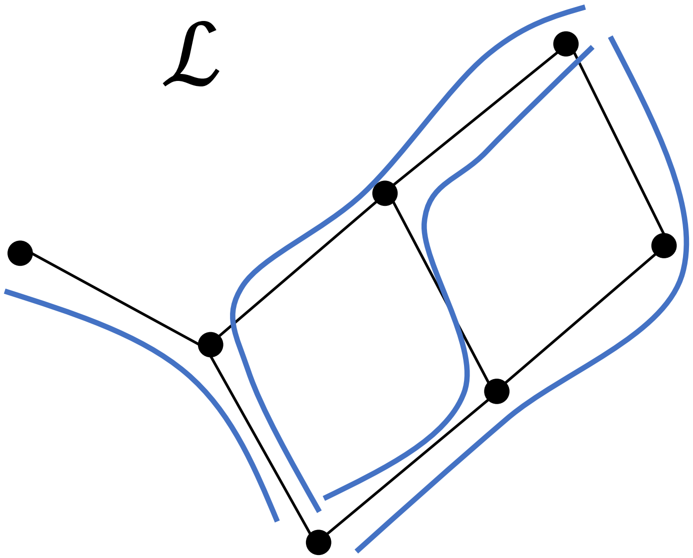
$K(\mathcal{M}) \coloneqq \{ \ \sum_i \mu_i p_i \ | \ p_0 \prec p_1 \prec \cdots \prec p_n, \ \ p_i \in \mathcal{M} \ \}$ = the order complex of $\mathcal{M}$

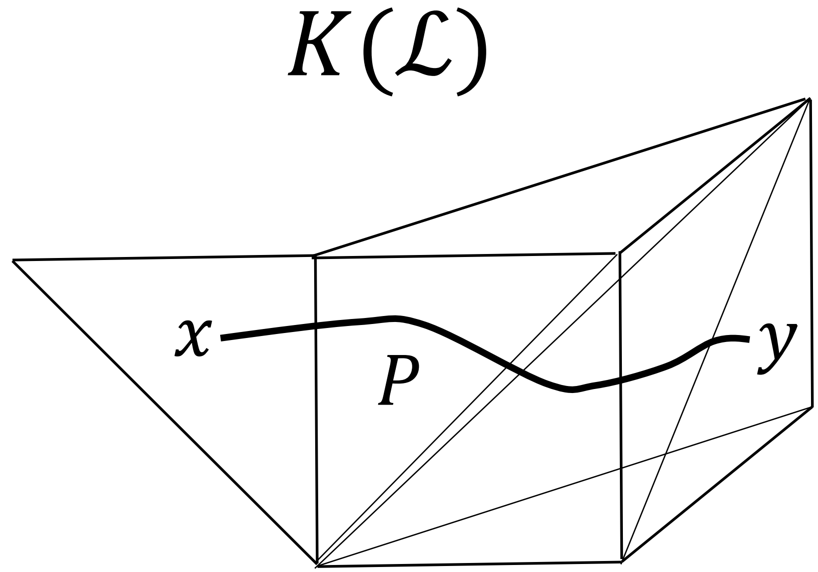
metrized by filling *orthoscheme* to each simplex:

$$\sum_i \mu_i p_i \longmapsto \sum_i \mu_i (\overbrace{1, \dots, 1}^{i}, 0, \dots, 0)$$



Orthoscheme = conv $\{(0,0,\dots,0), (1,0,\dots,0), (1,1,0,\dots,0), \dots, (1,1,1,\dots,1)\}$





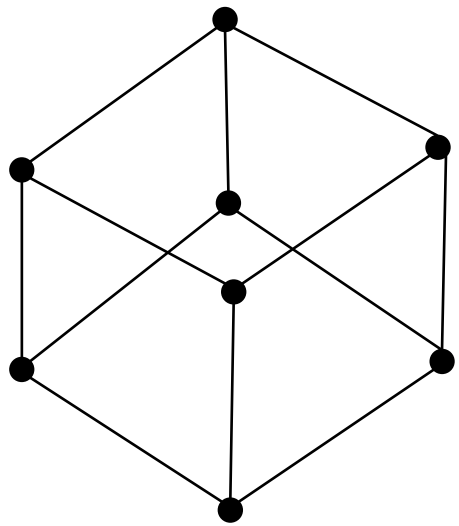
The length $d(P)$ of a path $P: [0,1] \to K(\mathcal{L})$ is well-defined

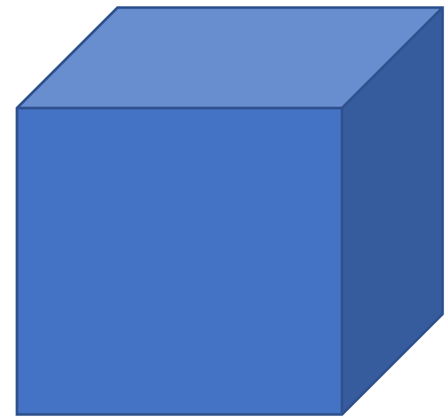
$$d(x, y) := \inf\{d(P) \mid P: (x, y)\text{–path}\}$$

$\to (K(\mathcal{L}), d)$ is a geodesic metric space

Question [Brady-McCammond 2012]:
Which is the class of posets $\mathcal{L}$ having CAT(0) $K(\mathcal{L})$?

Boolean Lattice





$\mathcal{L} = \{0,1\}^n$

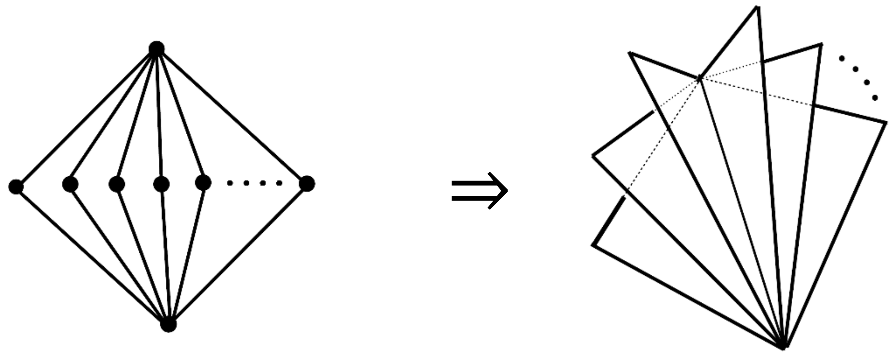
$K(\mathcal{L}) \simeq [0,1]^n$

$\rightarrow$ CAT(0)

$\mathcal{L}$: distributive lattice → $K(\mathcal{L})$: order polytope, CAT(0)

Thm [Haettel-Kielak-Schwer 2015, Chalopin-Chepoi-H-Osajda 2020]

$\mathcal{M}$: modular lattice $\Rightarrow K(\mathcal{M})$ is CAT(0)



$\mathcal{M}$: modular lattice of subspaces of $\mathbb{K}^2$

$K(\mathcal{M})$

Thm [H.2018] $\mathcal{M}$: modular lattice

$f: \mathcal{M} \to \mathbb{R}$ is submodular $\Leftrightarrow$ $\bar{f}: K(\mathcal{M}) \to \mathbb{R}$ is convex

$$\bar{f}(x) := \sum_i \mu_i f(p_i) \quad \left(x = \sum_i \mu_i p_i \, ; \, p_0 \prec p_1 \prec \cdots \prec p_n, \right)$$

Classical Lovász extension: $f: \{0,1\}^n \to \mathbb{R}$ is submodular $\Leftrightarrow \bar{f}: [0,1]^n \to \mathbb{R}$ is convex

$$\text{Min.}\quad -\dim X - \dim Y + C\textstyle\sum_i \operatorname{rank} A_i|_{X\times Y}$$

$$\text{s.t.}\quad (X,Y) \in \mathcal{L}\times\mathcal{L}^*$$

$\Downarrow$ Lovász extension

$$\text{Min.}\quad -\overline{\dim}(x) - \overline{\dim}\,(y) + C\textstyle\sum_i \overline{\operatorname{rank} A_i}(x,y)$$

$$\text{s.t.}\quad (x,y) \in K(\mathcal{L}\times\mathcal{L}^*) = K(\mathcal{L})\times K(\mathcal{L}^*)$$

→ convex optimization on CAT(0) space $K(\mathcal{L}\times\mathcal{L}^*)$

but we have no powerful method such as ellipsoid method

Feature: the objective is a sum of convex functions

→ splitting proximal point algorithm

Splitting Proximal Point Algorithm Bačák 2014

(K, d): complete CAT(0) space, $f_1, f_2, \dots, f_m$: convex func on K

Goal: Minimize $f = \sum_{i=1}^{m} f_i$

SPPA: Iterate

$$z^{k+1} \longleftarrow \operatorname{argmin}_{z \in K} \; f_{k \bmod m}(z) + \frac{1}{\lambda_k} d(z, z^k)^2$$

Thm (Ohta-Pálfia 2015)

f_i: L-Lipschitz, f: ε-strongly convex, $\lambda_k := \frac{1}{2\varepsilon(k+1)}$

$$\Rightarrow \qquad f(z^k) - \min f \le \frac{\mathrm{poly}(L, m, \varepsilon^{-1}, \mathrm{diam}\, K)}{\sqrt{k}}$$

$$\text{Min.} \quad -\overline{\dim}(x) - \overline{\dim}\,(y) + C\sum_i \overline{\operatorname{rank} A_i}(x,y)$$

+ Perturbation term for ε-strong convexity

$$\text{s.t.} \quad (x,y) \in K(\mathcal{L}\times\mathcal{L}^*) = K(\mathcal{L}) \times K(\mathcal{L}^*)$$

- Apply SPPA with:

  $f_0(x,y) \coloneqq -\overline{\dim}(x) - \overline{\dim}(y)$

  $f_i(x,y) \coloneqq C\,\overline{\operatorname{rank} A_i}(x,y) \quad (i = 1,2,\dots,m)$

- After $k = \operatorname{poly}(n,m)$ iterations, we have $f\left(x^k, y^k\right) - \mathrm{opt} < 1/2$

- By integrality of f, an optimum (X^*, Y^*) exists

  in the support of $\left(x^k, y^k\right) = \sum_i \mu_i (X_i, Y_i)$

Main technical contribution

Thm: SPAA is implementable in polytime, i.e.,

$$\text{Min. } f_i(x,y) + \frac{1}{\lambda} d((x,y),(x_0,y_0))^2$$

$$\text{s.t. } (x,y) \in K(\mathcal{L} \times \mathcal{L}^*)$$

where $f_i(x,y) = -\overline{\dim}(x) - \overline{\dim}(y)$ or

$f_i(x,y) = C\,\overline{\operatorname{rank} A_i}(x,y)$

is solvable in polytime.

Proof: Lattice theoretic / building theoretic argument

- An optimum exists in an *apartment* = the subcomplex of a Boolean sublattice including the supports of $x_0, y_0, x_0^{\perp_{A_i}}, y_0^{\perp_{A_i}}$
- In the apartment $\approx [0,1]^{2n}$, the problem is an easy convex quadratic program

Good / bad points

- Conceptually simple, described uniformly in an arbitrary field
- Applications to other problems ?

- Polytime, but very slow:

  ← Less use of characteristics of the objective function

  ← Undevelopment of optimization theory in CAT(0) spaces

- Bit complexity problem for $\mathbb{K} = \mathbb{Q}$:

  ← Polynomial number of vector-space operations $X \cap X', X + X'$

  can cause exponential bit explosion of bases

Reduction of nc-singularity over $\mathbb{Q}$ to that over $\mathbb{F}_p$ by p-adic valuation v_p & discrete convex optimization on Euclidean building

$$\begin{aligned} \text{Min.} \quad & v_p \det P + v_p \det Q \\ \text{s.t.} \quad & v_p(PA_kQ)_{ij} \geq 0 \quad (\forall ij, k), \\ & P, Q \in GL_n(\mathbb{Q}) \end{aligned}$$

- Nc-singular $\Leftrightarrow$ optimal value $= -\infty$

$$z = \underbrace{1011\,\overbrace{000}^{v_2(z)}}_{bitsize}$$

- opt $> -\infty \Rightarrow$ opt $> -\text{poly (input)}$ ← $v_p(|z|) \leq$ bit-size of z for $0 \neq z \in \mathbb{Z}$
- *L-convex function* minimization on Euclidean building for $GL_n(\mathbb{Q})$
- Descent algorithm [H.2019], where each step is nc-rank computation over $\mathbb{F}_p$

Application to related problems

(p, q)-scalability of $\{A_k\}$ by triangular matrices [Franks 2018]

$$\text{Max.} \quad \sum_i (p_i - p_{i+1}) \dim E_i \cap X + \sum_i (q_i - q_{i+1}) \dim F_i \cap Y$$

E_i, F_i:standard flags

$$\text{s.t.} \quad A_k(X, Y) = \{0\} \quad (\forall k)$$

X, Y: vector subspaces

Brascamp-Lieb separation for BL-data (B, p) [Garg et al. 2018]

$$\text{Max.} \quad \dim X + \sum_i p_i \dim Y_i$$

$$\text{s.t.} \quad B_i(X, Y_i) = \{0\} \quad (\forall i)$$

X, Y_i: vector subspaces

Our approach is applicable, but it only gives pseudo-polynomial complexity

I don’t know whether p-adic reduction can work

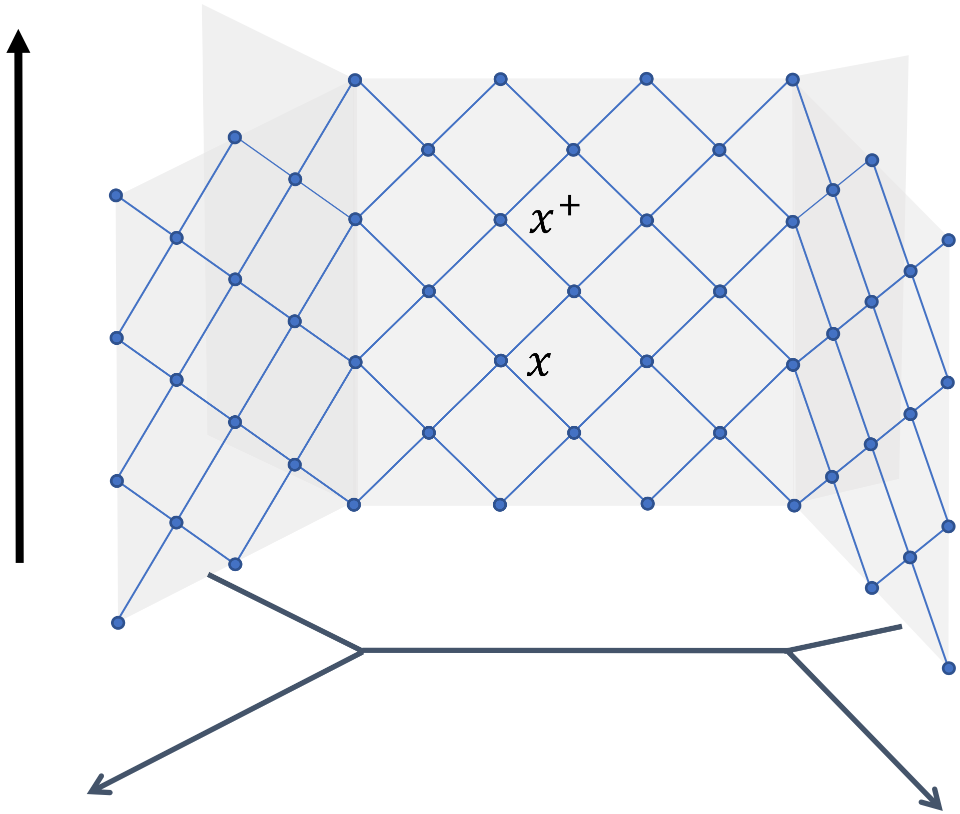
Reference

M. Hamada and H. Hirai: Computing the nc-rank via discrete convex optimization on CAT(0) spaces, *SIAM Journal on Applied Geometry and Algebra (SIAGA),* to appear.

Papers and slides are available at

http://www.misojiro.t.u-tokyo.ac.jp/~hirai/

Thank you for your attention !



L-convex function on Euclidean building (of type A) [H. 2019]

$$f(x^+) = f(x) + \text{const}$$

$$f(x) + f(y) \geq f(x \wedge y) + f(x \vee y)$$

extends the one in *Discrete Convex Analysis* on $\mathbb{Z}^n$ [Murota 1998]

→ *Discrete Convex Analysis beyond* $\mathbb{Z}^n$ [H. 2013～]