Log-barrier interior point methods are not strongly polynomial

Michael Joswig

TU Berlin

Simons Institute, 20 Oct 2017

joint w/ Xavier Allamigeon
Pascal Benchimol
Stéphane Gaubert

Main Results

Theorem (ABGJ 2017+)

There is a family, $\mathbf{LW}_r(t)$, of linear programs in $2r$ variables with $3r+1$ constraints, depending on $t > 1$, such the number of iterations of any primal-dual path-following interior point algorithm with a log-barrier function which iterates in the wide neighborhood of the central path is exponential in r for $t \gg 0$.

Theorem (ABGJ 2014+)

On the same family of LPs the total curvature of the central path is in $\Omega(2^r)$ for $t \gg 0$.

Ridiculously Abbreviated History

Algorithms

- Karmarkar 1984: polynomial time interior point algorithm
- Renegar 1988: $O(\sqrt{m+n}\,L)$
  - where $L =$ total bit size of input
- wide neighborhood methods:
  - short/long step: Kojima, Mizuno & Yoshise 1989, Monteiro & Adler 1989
  - predictor-corrector: Mizuno, Todd & Ye 1993, Vavasis & Ye 1996

Geometry

- Bayer & Lagarias 1989; Dedieu & Shub 2005; Dedieu, Malajovich and Shub 2005: curvature of central path
- Deza, Terlaky & Zinchenko 2009: redundant Klee–Minty cube
  - continuous Hirsch conjecture

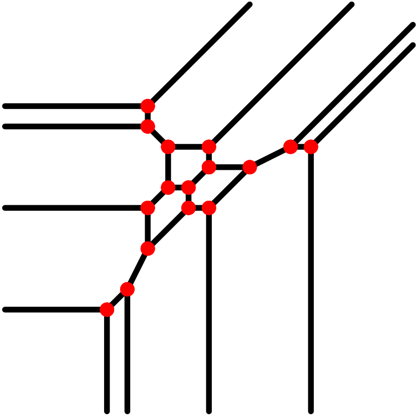
The Linear Programs $\mathbf{LW}_r(t)\mathbf{LW}_r^{\epsilon}(t)$ ...

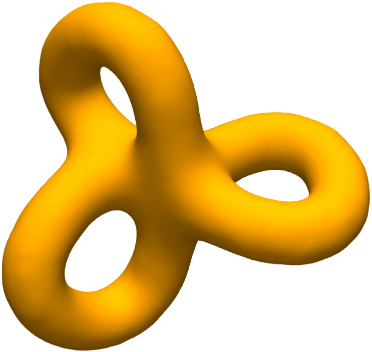
$$
\begin{aligned}
\text{minimize} \quad & x_1 \\
\text{subject to} \quad & x_1 \le t^2 \\
& x_2 \le t \\
& \left.\begin{aligned} & x_{2j+1} \le t\, x_{2j-1}\,,\; x_{2j+1} \le t\, x_{2j} \\ & x_{2j+2} \le t^{1-1/2^j}(x_{2j-1} + x_{2j}) \end{aligned}\right] \; 1 \le j < r \\
& x_{2r-1} \ge 0\,,\; x_{2r} \ge 0\epsilon
\end{aligned}
$$

for $r \ge 1$ and $t \gg 0$
and $1 \gg \epsilon \ge 0$

... have long and winding central paths.

"Piecewise linear shadows of classical varieties"





$$t^8(x^4+y^4+z^4)+t^4(x^3y+xz^3+y^3z)+t^2(x^3z+xy^3+yz^3)$$
$$+t(x^2y^2+x^2z^2+y^2z^2)+(x^2yz+xy^2z+xyz^2)$$

where $t \gg 0$

Tropical Arithmetic

tropical semi-ring: $\mathbb{T} = \mathbb{T}(\mathbb{R}) = (\mathbb{R} \cup \{-\infty\}, \oplus, \odot)$ where

$$x \oplus y \ := \ \max(x, y) \quad \text{and} \quad x \odot y \ := \ x + y$$

- absolutely convergent (generalized) Puiseux series with real coefficients

$$\mathbb{R}_{\text{conv}}\{\{t\}\} \ = \ \{\underbrace{c_{\alpha_1} t^{\alpha_1} + c_{\alpha_2} t^{\alpha_2} + \cdots}_{\gamma(t)}\} \cup \{0\}$$

  such that $\alpha_1 > \alpha_2 > \cdots$ strictly descending sequence of reals (finite or unbounded), $c_{\alpha_i} \in \mathbb{R} - \{0\}$, absolutely convergent for $t \gg 0$
  $\rightsquigarrow$ real closed — Dries & Speissegger 1998

- valuation map $\text{ord}(\gamma(t)) = \alpha_1$ and $\text{ord}(0) = -\infty$

$$\text{ord}(\gamma(t) + \delta(t)) \ \leq \ = \ \max(\text{ord}(\gamma(t)), \text{ord}(\delta(t))) \, \text{ord}(\gamma(t)) \oplus \text{ord}(\delta(t))$$
$$\text{ord}(\gamma(t) \cdot \delta(t)) \ = \ \text{ord}(\gamma(t)) + \text{ord}(\delta(t)) \, \text{ord}(\gamma(t)) \odot \text{ord}(\delta(t))$$

Tropicalization

The polynomial

$$f \; = \; \gamma(t) x_1^{u_1} x_2^{u_2} \dots x_d^{u_d} + \delta(t) x_1^{v_1} x_2^{v_2} \dots x_d^{v_d} + \dots$$

gives rise to the tropicalization

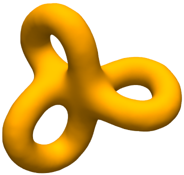
$$\begin{aligned} F \; = \; \mathrm{trop}(f) \; := \; & \mathrm{ord}(\gamma(t)) \odot x_1^{\odot u_1} \odot x_2^{\odot u_2} \odot \cdots \odot x_d^{\odot u_d} \\ & \oplus \mathrm{ord}(\delta(t)) \odot x_1^{\odot v_1} \odot x_2^{\odot v_2} \odot \cdots \odot x_d^{\odot v_d} \oplus \dots , \end{aligned}$$

where $\mathrm{ord}(\gamma(t)) =$ highest t-exponent

Example

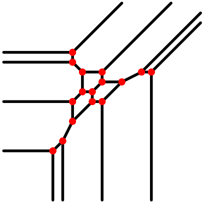
$$\begin{array}{llllll} f & = & x^3 - & (t^3 + 2t + 1)x^2 + & (2t^4 + t^3 + 2t)x - & 2t^4 \\ F & = & x^{\odot 3} \oplus & 3 \odot x^{\odot 2} \;\; \oplus & 4 \odot x \;\; \oplus & 4 \\ & = \max(& 3x \quad , & 3 + 2x \quad , & 4 + x \quad , & 4 \;) \end{array}$$

Main Theorem of Tropical Geometry



Theorem (Kapranov 2002)

For $f \in \mathbb{C}\{\{t\}\}[x_1, x_2, \ldots, x_d]$ the tropical hypersurface $\mathcal{T}(F)$ coincides with $\text{ord}(V(f))$.



Definition

F vanishes if maximum attained at least twice

Example

$f = x^3 - (t^3 + 2t + 1)x^2 + (2t^4 + t^3 + 2t)x - 2t^4$ vanishes at $x = 2t$

$F = \max(3x, \boxed{3+2x}, \boxed{4+x}, 4)$ vanishes at $x = 1 = \text{ord}(2t)$

Example: The Linear Assignment Problem

Problem

Given 4 football players and 4 positions, what is the best formation?

$$A = \begin{pmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & 3 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

- assignment = choice of coefficients, one per column/row

$$\text{best} = \max_{\omega \in \mathrm{Sym}(4)} a_{1,\omega(1)} + a_{2,\omega(2)} + a_{3,\omega(3)} + a_{4,\omega(4)}$$

$$= \bigoplus_{\omega \in \mathrm{Sym}(4)} a_{1,\omega(1)} \odot a_{2,\omega(2)} \odot a_{3,\omega(3)} \odot a_{4,\omega(4)}$$

Definition (tropical determinant)

tdet = trop(det)

Linear Programming via Interior Point Method

Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$, $\mu > 0$.

primal linear program:

assume bounded w/ non-empty interior

$$\begin{aligned} &\text{minimize} && c^\top x \\ &\text{subject to} && Ax \leq b,\ x \geq 0,\ x \in \mathbb{R}^n \end{aligned} \qquad \text{LP}(A, b, c)$$

dual linear program:

$$\begin{aligned} &\text{maximize} && -b^\top y \\ &\text{subject to} && -A^\top y \leq c,\ y \geq 0,\ y \in \mathbb{R}^m \end{aligned}$$

associated logarithmic barrier problem:

$$\begin{aligned} &\text{minimize} && \frac{c^\top x}{\mu} - \sum_{j=1}^{n} \log(x_j) - \sum_{i=1}^{m} \log(w_i) \\ &\text{subject to} && Ax + w = b,\ x > 0, w > 0 \end{aligned}$$

A System of Polynomial Equations

logarithmic barrier problem

$$\begin{array}{ll} \text{minimize} & \dfrac{c^\top x}{\mu} - \displaystyle\sum_{j=1}^{n} \log(x_j) - \sum_{i=1}^{m} \log(w_i) \\ \text{subject to} & Ax + w = b,\ x > 0, w > 0 \end{array}$$

for $\mu > 0$ has unique optimal solution (x^μ, w^μ) chacterized by

$$\begin{aligned} Ax + w &= b \\ -A^\top y + s &= c \\ w_i y_i &= \mu \quad \text{for all } i \in [m] \\ x_j s_j &= \mu \quad \text{for all } j \in [n] \\ x, w, y, s &> 0 \end{aligned}$$

That is, there uniquely exist y^μ and s^μ such that $(x^\mu, w^\mu, y^\mu, s^\mu)$ is a solution ...

The Central Path and the Central Curve

Definition

The central path is the image of the map

$$\mathcal{C}_{A,b,c} : \mathbb{R}_{>0} \to \mathbb{R}^{2m+2n}, \quad \mu \mapsto (x^\mu, w^\mu, y^\mu, s^\mu) .$$

- primal central path = projection onto x-coordinates
- dual central path = projection onto y-coordinates

Observation

The equality constraints in the log-barrier problem define a real algebraic curve, the central curve, which is the Zariski closure of the central path.

The Wide Neighborhood

Let $z = (x, w, s, y) \in \mathbb{R}^{2n+2m}$.
For duality measure $\bar{\mu}(z) := \frac{1}{n+m}(\langle x, s\rangle + \langle w, y\rangle)$ we have

$$z = \mathcal{C}(\mu) \iff \begin{pmatrix} xs \\ wy \end{pmatrix} = \bar{\mu}(z)e$$

Yields a first neighborhood (e.g., for ℓ_2-norm)

$$\mathcal{N}_\theta := \left\{ z \in \mathcal{F}^\circ \colon \left\| \begin{pmatrix} xs \\ wy \end{pmatrix} - \bar{\mu}(z)e \right\| \le \theta\bar{\mu}(z) \right\}$$

for some real precision parameter $\theta \in (0, 1)$.
This is replaced by the wide neigborhood

$$\mathcal{N}_\theta^{-\infty}(\mu) := \left\{ z \in \mathcal{F}^\circ \colon \begin{pmatrix} xs \\ wy \end{pmatrix} \ge (1-\theta)\bar{\mu}(z)e \right\}$$

for the one-sided ℓ_∞-norm $\max(0, \max_k(-v_k))$.

Maslov Dequantization of Central Paths

For $\boldsymbol{A} \in \mathbb{K}^{m \times n}$, $\boldsymbol{b} \in \mathbb{K}^m$ and $\boldsymbol{c} \in \mathbb{K}^n$ assume

$$\boldsymbol{\mathcal{P}} = \{\boldsymbol{x} \in \mathbb{K}^n \mid \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \boldsymbol{x} \geq 0\}$$

bounded with non-empty interior. Not necessarily compact!

- $\mathbb{K} = \mathbb{R}_{\mathsf{conv}}\{\{t\}\}$ absolutely convergent generalized Puiseux series
- for $t \gg 0$ real linear programs $\mathrm{LP}(\boldsymbol{A}(t), \boldsymbol{b}(t), \boldsymbol{c}(t))$ well defined
- $\mathcal{C}(t, \lambda) = \mathcal{C}_{\boldsymbol{A}(t), \boldsymbol{b}(t), \boldsymbol{c}(t)}(t^\lambda)$ real central path

Definition

$$\mathcal{C}^{\mathrm{trop}} : \lambda \mapsto \lim_{t \to +\infty} \log_t \mathcal{C}(t, \lambda) \qquad \text{tropical central path}$$

Proposition (ABGJ 2017+)

The family of maps $(\log_t \mathcal{C}(t, \cdot))_t$ converges uniformly on any closed interval $[a, b] \subset \mathbb{R}$ to the tropical central path $\mathcal{C}^{\mathrm{trop}}$.

Tropicalizing a System of Linear Inequalities

Consider the Puiseux polyhedron $\boldsymbol{\mathcal{P}} \subset \mathbb{K}^2$ defined by:

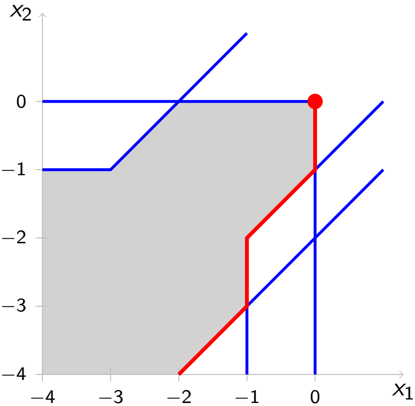
$$\begin{aligned} \boldsymbol{x}_1 + \boldsymbol{x}_2 &\leq 2 \\ t\boldsymbol{x}_1 &\leq 1 + t^2\boldsymbol{x}_2 \\ t\boldsymbol{x}_2 &\leq 1 + t^3\boldsymbol{x}_1 \\ \boldsymbol{x}_1 &\leq t^2\boldsymbol{x}_2 \\ \boldsymbol{x}_1, \boldsymbol{x}_2 &\geq 0 \ . \end{aligned} \tag{1}$$

Then the set $\operatorname{ord}(\boldsymbol{\mathcal{P}})$ is described by the tropical linear inequalities:

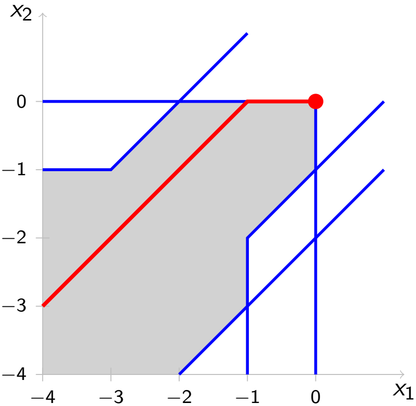
$$\begin{aligned} \max(x_1, x_2) &\leq 0 \\ 1 + x_1 &\leq \max(0, 2 + x_2) \\ 1 + x_2 &\leq \max(0, 3 + x_1) \\ x_1 &\leq 2 + x_2 \ . \end{aligned} \tag{2}$$

. . . and Two of Its Primal Tropical Central Paths

- tropical central path = ord(Puiseux central path)



$\min \boldsymbol{x}_1$



$\min t\boldsymbol{x}_1 + \boldsymbol{x}_2$

Maslov Dequantization of Central Paths

Recall the claim:

Proposition (ABGJ 2017+)

The family of maps $(\log_t \mathcal{C}(t,\cdot))_t$ converges uniformly on any closed interval $[a,b] \subset \mathbb{R}$ to the tropical central path $\mathcal{C}^{\mathrm{trop}}$.

Proof of Dequantization Theorem

$z_t :=$ function $\lambda \mapsto \log_t \mathcal{C}(t, \lambda) \in \mathbb{R}^{2n+2m}$ $\qquad$ $z := \lim_{t\to\infty} z_t$ pointwise

Proof.

Fix $\epsilon > 0$ and choose partition $a = a_1 < a_2 < \cdots < a_k < a_{k+1} = b$ such that $a_{i+1} - a_i \leq \epsilon$ for all i. Pick $\lambda \in [a_i, a_{i+1}]$. Then

$$|z_t(\lambda) - z(\lambda)| \leq ?|z_t(\lambda) - z_t(a_i)|2\epsilon + |z_t(a_i) - z(a_i)| + |z(a_i) - z(\lambda)|\epsilon\,.$$

Can show:

$$|z_t(\lambda) - z_t(a_i)| \leq \log_t(2n + 2m) + \lambda - a_i \leq \log_t(2n + 2m) + \epsilon$$

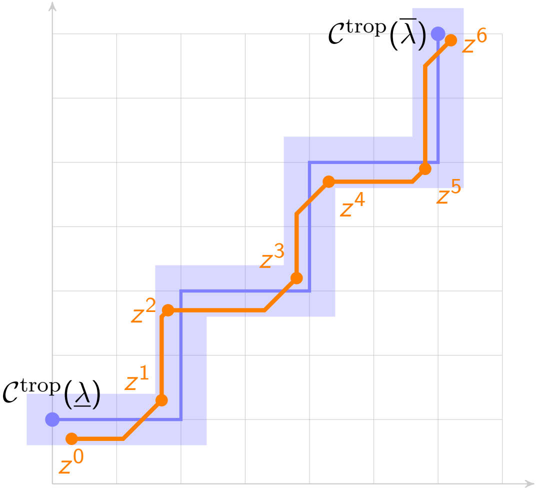
Thus, there exists t_ϵ with $|z_t(\lambda) - z_t(a_i)| \leq 2\epsilon$ for all $t \geq t_\epsilon$.
Can also show:

$$|z(\lambda) - z(a_i)| \leq \lambda - a_i \leq \epsilon$$

Pointwise convergence takes care of final term. □

Tubular Neighborhood Controls Iteration Complexity

- number of tropical segments required to approximate tropical central path bounded from below



Recall: $\mathbf{LW}_r(t)\mathbf{LW}_r^{\epsilon}(t)$

$$
\begin{array}{rl}
\text{minimize} & x_1 \\
\text{subject to} & x_1 \leq t^2 \\
& x_2 \leq t \\
& \left.\begin{array}{l} x_{2j+1} \leq t\, x_{2j-1}\,,\ x_{2j+1} \leq t\, x_{2j} \\ x_{2j+2} \leq t^{1-1/2^j}(x_{2j-1} + x_{2j}) \end{array}\right] \ 1 \leq j < r \\
& x_{2r-1} \geq 0\,,\ x_{2r} \geq 0\epsilon
\end{array}
$$

for $r \geq 1$ and $t \gg 0$
and $1 \gg \epsilon \geq 0$

An Explicit Bound for t

Theorem (ABGJ 2017+)

Let $0 < \theta < 1$, and suppose that

$$t > \left(\frac{\big((10r-1)!\big)^8}{1-\theta} \right)^{2^{r+2}} .$$

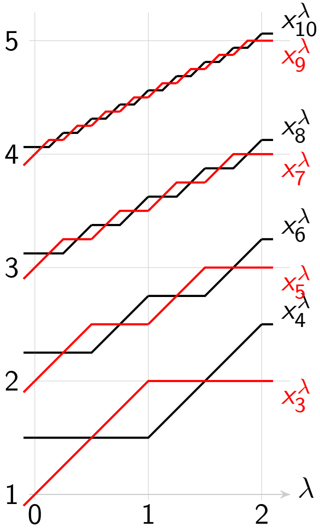
Then, every polygonal curve $[z^0, z^1] \cup [z^1, z^2] \cup \cdots \cup [z^{p-1}, z^p]$ contained in the neighborhood $\mathcal{N}_{\theta,t}^{-\infty}$ of the primal-dual central path of $\mathbf{LW}_r^{=}(t)$, with $\bar{\mu}(z^0) \leq 1$ and $\bar{\mu}(z^p) \geq t^2$, contains at least 2^{r-1} segments.

duality measure

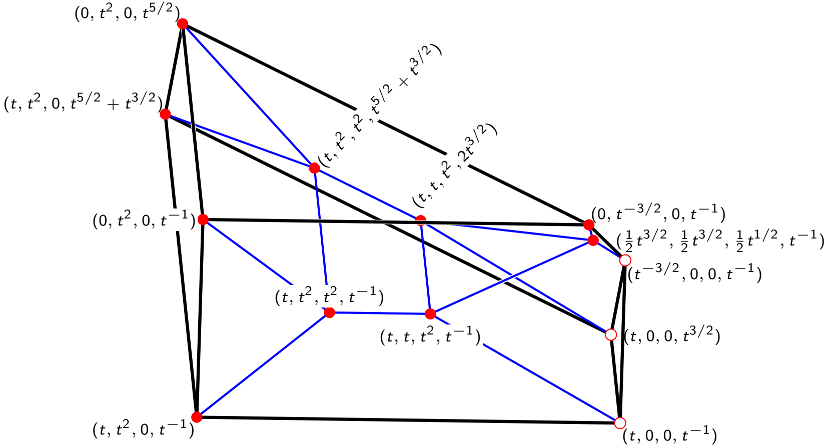
$$\bar{\mu}(z) \coloneqq \frac{1}{n+m}\big(\langle x, s\rangle + \langle w, y\rangle\big)$$

The Tropical Central Paths of the Counter-Examples

- the x-components of the primal tropical central path of $\mathbf{LW}_r$ for $r \geq 5$ and $0 \leq \lambda \leq 2$
- lifting a construction by Bezem, Nieuwenhuis and Rodríguez-Carbonell 2008



Schlegel Diagram of $\mathbf{LW}_2(2)$, perturbed to simplicity



$n = 4,\ m = 7$

Conclusion

- tropical geometry is useful for getting insight about intricate details in (linear) optimization
- sheds new light on the interior point method as well as on the simplex method

Allamigeon, Benchimol, Gaubert & J.:

1. *Tropicalizing the simplex algorithm*, SIAM J. Discrete Math. **29** (2015)
2. *Combinatorial simplex algorithms can solve mean payoff games*, SIAM J. Opt. **24** (2014)
3. *Long and winding central paths*, `arXiv:1405.4161`
4. *Log-barrier interior point methods are not strongly polynomial*, to appear in SIAM J. Appl. Alg. Geo., `arXiv:1708.01544`

Uniform Convergence

$$\begin{aligned} \delta_{\mathrm{F}}(x,y) &:= \max(0, \max_k(y_k - x_k)) && \text{Funk metric} \\ d_\infty(x,y) &:= \max(\delta_{\mathrm{F}}(x,y), \delta_{\mathrm{F}}(y,x)) && \text{symmetrized Funk} \\ d_{\mathrm{H}}(x,y) &:= \delta_{\mathrm{F}}(x,y) + \delta_{\mathrm{F}}(y,x) && \text{Hilbert's projective metric} \\ \delta(t) &:= 2 d_{\mathrm{H}}(\log_t \boldsymbol{\mathcal{F}}(t), \mathcal{F}) && \text{deviation of feasible regions} \end{aligned}$$

Theorem (ABGJ 2017+)

For all $t > t_0$ and $\mu > 0$ we have

$$d_\infty\left(\log_t \mathcal{N}_{\theta,t}^{-\infty}(\mu), \mathcal{C}^{\mathrm{trop}}(\log_t \mu)\right) \leq \log_t\left(\frac{N}{1-\theta}\right) + \delta(t)\,.$$

Metric Estimate For Maslov Dequantization of Polyhedra

Theorem (ABGJ 2017+)

Let $\mathcal{P} \subset \mathbb{K}_+^d$ *be a polyhedron of the form* $\{\boldsymbol{x} \in \mathbb{K}^d : \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\}$ *where* $\boldsymbol{A}$ *and* $\boldsymbol{b}$ *are monomial. Let* η_0 *be the minimum of the quantities* $\eta(\boldsymbol{M})$ *where* $\boldsymbol{M}$ *is a square submatrix of* $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{b} & 0 \\ e^\top & 0 & 1 \end{pmatrix}$ *of order* d*.*

Then, for all $t \geq (d!)^{1/\eta_0}$*, we have:*

$$d_{\mathrm{H}}(\log_t \mathcal{P}(t), \operatorname{ord}(\mathcal{P})) \leq \log_t\big((d+1)^2(d!)^4\big)\,.$$

$$\eta(\boldsymbol{M}) \coloneqq \min\Big\{\eta \colon \sigma, \tau \in \operatorname{Sym}(d),\ \eta = \sum_{i=1}^{d} \alpha_{i\sigma(i)} - \sum_{i=1}^{d} \alpha_{i\tau(i)} > 0\Big\}$$

Tubular Neighborhood

Theorem (ABGJ 2017+)

For $0 < \theta < 1$ *suppose that* $t > t_0$ *satisfies*

$$\log_t\Big(\frac{2N}{1-\theta}\Big) + \delta(t) < \epsilon_0\big([\underline{\lambda}, \overline{\lambda}]\big) \, .$$

Then, every polygonal curve $[z^0, z^1] \cup [z^1, z^2] \cup \cdots \cup [z^{p-1}, z^p]$ *contained in the neighborhood* $\mathcal{N}_{\theta,t}^{-\infty}$*, with* $\bar{\mu}(z^0) \leq t^{\underline{\lambda}}$ *and* $\bar{\mu}(z^p) \geq t^{\overline{\lambda}}$*, contains at least* $\gamma\big([\overline{\lambda}, \underline{\lambda}]\big)$ *segments.*

Geometric Characterization of Tropical Central Path

Fix $\boldsymbol{\mu} \in \mathbb{K}$ positive.

$$
\begin{aligned}
(\boldsymbol{x}^{\boldsymbol{\mu}}, \boldsymbol{w}^{\boldsymbol{\mu}}) &= \text{corresponding point on primal central path of LP}(\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}) \\
\boldsymbol{\nu} &= \text{that LP's optimal value} \\
\boldsymbol{\mathcal{P}}^{\boldsymbol{\mu}} &= \{(\boldsymbol{x}, \boldsymbol{w}) \in \mathbb{K}_+^{n+m} \mid \boldsymbol{A}\boldsymbol{x} + \boldsymbol{w} = \boldsymbol{b},\ \boldsymbol{c}\boldsymbol{x} \leq \boldsymbol{\nu} + (n+m)\boldsymbol{\mu}\}
\end{aligned}
$$

Theorem (ABGJ 2014+)

Then $\operatorname{ord}(\boldsymbol{x}^{\boldsymbol{\mu}}, \boldsymbol{w}^{\boldsymbol{\mu}})$ *equals tropical barycenter of* $\operatorname{ord}(\boldsymbol{\mathcal{P}}^{\boldsymbol{\mu}})$.