Certifying the Radius of Positive Semidefiniteness Via Our ArtinProver Package

Erich L. Kaltofen
North Carolina State University
`google->kaltofen`

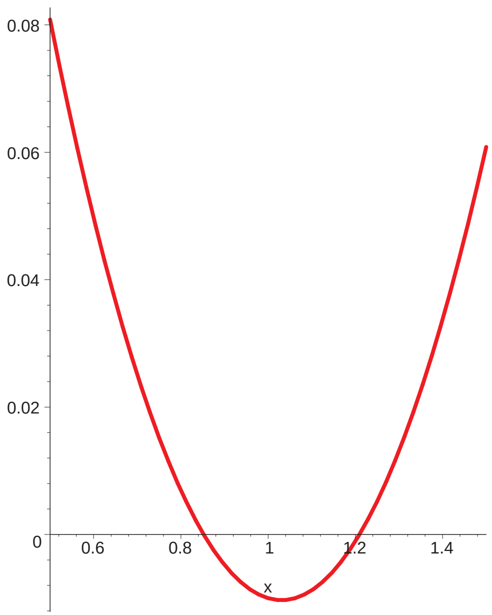
Joint with Sharon E. Hutton and Lihong Zhi

Inequalities with floating point scalars

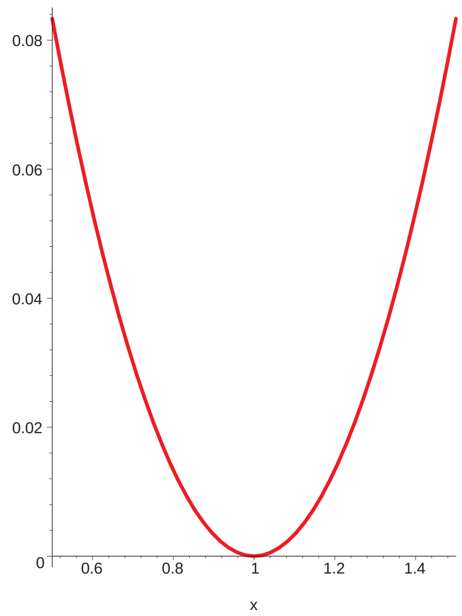
$$\forall x \in \mathbb{R}:\ 0.34x^2 - 0.66x + 0.33 > 0.009 \quad ???$$

Inequalities with floating point scalars

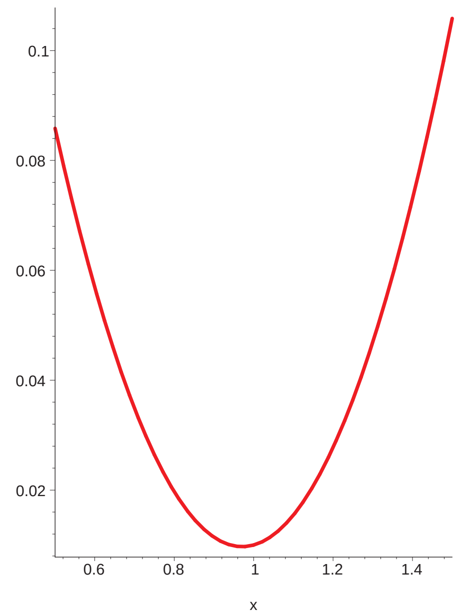
$$\forall x \in \mathbb{R} :\ 0.34x^2 - 0.66x + 0.33 > 0.009 \quad ???$$



$$\left(\frac{1}{3} - \frac{1}{100}\right)x^2 - \frac{2}{3}x + \frac{1}{3}$$



$$\frac{1}{3}x^2 - \frac{2}{3}x + \frac{1}{3}$$



$$\left(\frac{1}{3} + \frac{1}{100}\right)x^2 - \frac{2}{3}x + \frac{1}{3}$$

Radius of positive semidefinitenes (unconstrained coeff's)

$$\rho_2(f) = \inf_{\tilde{f} \in \mathbb{R}[x_1,\ldots,x_n]} \| f - \tilde{f} \|_2^2 \quad \text{(coeff. vector 2-norm)}$$

$$\text{s. t. } \exists \alpha \in \mathbb{R}^n \colon \ \tilde{f}(\alpha) = 0,$$

$$\deg(\tilde{f}) \le \deg(f). \quad \text{(any degree notion)}$$

Note: $\rho_2(\, x^2 + 1\,) = 1, \quad \tilde{f} = x^2$

Note: $\rho_2(\, x^2 + (1 - xy)^2\,) = 0$ as $\varepsilon^2 + (1 - \varepsilon \frac{1}{\varepsilon})^2 - \varepsilon^2 = 0$

Main Theorem [Hitz, Hutton, Kaltofen, Karmarkar Lakshman, Sciabica, Ruatta, Szanto, Zhi]

Let $\alpha \in \mathbb{R}^n$: $\mathcal{N}_2^{[f]}(\alpha) = \inf_{\tilde{f} \in \mathbb{R}[x_1,\ldots,x_n]} \|f - \tilde{f}\|_2^2$

s. t. $\tilde{f}(\alpha) = 0,$

$\deg(\tilde{f}) \leq \deg(f)$

$$\boxed{= \frac{f(\alpha)^2}{\|\tau\|_2^2}},$$

where $\tau = [1, \alpha_1, \ldots, \alpha_n, \ldots, \alpha_1^{i_1} \alpha_2^{i_2} \cdots \alpha_n^{i_n}, \ldots]_{(i_1,\ldots,i_n) \leq \deg(f)}$

The coefficient vector $\vec{\tilde{f}}$ of the minimizer is $\boxed{\vec{\tilde{f}} = \vec{f} - \frac{\tau^T \vec{f}}{\|\tau\|_2^2} \tau}$

Main Theorem [Hitz, Hutton, Kaltofen, Karmarkar Lakshman, Sciabica, Ruatta, Szanto, Zhi]

Let $\alpha \in \mathbb{R}^n$: $\mathcal{N}_2^{[f]}(\alpha) = \inf_{\tilde{f} \in \mathbb{R}[x_1,\ldots,x_n]} \|f - \tilde{f}\|_2^2$

s. t. $\tilde{f}(\alpha) = 0,$

$\deg(\tilde{f}) \le \deg(f)$

$$\boxed{= \frac{f(\alpha)^2}{\|\tau\|_2^2}},$$

where $\tau = [1, \alpha_1, \ldots, \alpha_n, \ldots, \alpha_1^{i_1} \alpha_2^{i_2} \cdots \alpha_n^{i_n}, \ldots]_{(i_1,\ldots,i_n) \le \deg(f)}$

The coefficient vector $\vec{\tilde{f}}$ of the minimizer is $\boxed{\vec{\tilde{f}} = \vec{f} - \frac{\tau^T \vec{f}}{\|\tau\|_2^2} \tau}$

Corollary:

$$\rho_2(f) = \inf_{(\alpha_1,\ldots,\alpha_n) \in \mathbb{R}^n} \frac{f(\alpha_1, \ldots, \alpha_n)^2}{\sum_{(i_1,\ldots,i_n) \le \deg(f)} \alpha_1^{2i_1} \cdots \alpha_n^{2i_n}}$$

Our contributions:

- New proof by Lagrangian multipliers
  $\longrightarrow$ customized formulas for equality constraints on coeff's
  $\longrightarrow$ can have inequality constraints via
  Karush-Kuhn-Tucker conditions
  for fixed root α get linear program

Our contributions:

- New proof by Lagrangian multipliers
  $\longrightarrow$ customized formulas for equality constraints on coeff's
  $\longrightarrow$ can have inequality constraints via
  Karush-Kuhn-Tucker conditions
  for fixed root α get linear program

- SOS certificates for rational lower bound $\tilde{\rho}_2(f) < \rho_2(f)$
  (Seidenberg's problem with imprecise coefficients)

Example $f(x) = 1 - 2xy + x^2y^2 + x^2, \quad \alpha = (\varepsilon, \frac{1}{\varepsilon})$

$$\mathscr{N}_2(\varepsilon, \frac{1}{\varepsilon}) = \frac{\varepsilon^4}{\delta} < \frac{\varepsilon^4}{3},$$

$$\delta = 3 + 2\varepsilon^2 + \frac{2}{\varepsilon^2} + 2\varepsilon^4 + \frac{2}{\varepsilon^4} + \varepsilon^6 + \frac{1}{\varepsilon^6} + \varepsilon^8 + \frac{1}{\varepsilon^8},$$

$$\tilde{f}(x,y) = -\frac{\varepsilon^6}{\delta}x^4 - \frac{\varepsilon^4}{\delta}x^3y + (1 - \frac{\varepsilon^2}{\delta})x^2y^2 - \frac{1}{\delta}xy^3 - \frac{1}{\varepsilon^2\delta}y^4$$
$$- \frac{\varepsilon^5}{\delta}x^3 - \frac{\varepsilon^3}{\delta}x^2y - \frac{\varepsilon}{\delta}xy^2 - \frac{1}{\varepsilon\delta}y^3 + (1 - \frac{\varepsilon^4}{\delta})x^2$$
$$- (2 + \frac{\varepsilon^2}{\delta})xy - \frac{1}{\delta}y^2 - \frac{\varepsilon^3}{\delta}x - \frac{\varepsilon}{\delta}y + 1 - \frac{\varepsilon^2}{\delta}.$$

Example $f(x,y) = x^2 + y^2 + 1$

Nearest unconstrained solution $\tilde{f} = x^2 + 0\,xy + y^2 + 0\,y + 0$

Coefficient constraints of $\tilde{f}$: $\vec{\tilde{\tilde{f}}}_{1,1} = \vec{\tilde{\tilde{f}}}_{0,0}$ and $\vec{\tilde{\tilde{f}}}_{0,1} = 0$

Add constraints to Lagrangian and solve:

$$\frac{3\,\alpha^4 + 2\,\alpha^3\beta + 5\,\alpha^2\beta^2 + 3\,\alpha^2 + 2\,\alpha\beta + 2\,\alpha\beta^3 + 1 + 3\,\beta^4 + 2\,\beta^2}{2\,\alpha^2 + 2\,\alpha^4 + 2\,\beta^4 + \alpha^2\beta^2 + 2\,\alpha\beta + 1}$$

$$\neq \frac{f(\alpha,\beta)^2}{1 + \alpha^2 + \beta^2 + \alpha^2\beta^2 + \alpha^4 + \beta^4}$$

Example $f(x,y) = x^2 + y^2 + 1$

Nearest unconstrained solution $\tilde{f} = x^2 + 0\,xy + y^2 + 0\,y + 0$

Coefficient constraints of $\tilde{f}$: $\vec{\tilde{f}}_{1,1} = \vec{\tilde{f}}_{0,0}$ and $\vec{\tilde{f}}_{0,1} = 0$

Add constraints to Lagrangian and solve:

$$\frac{3\,\alpha^4 + 2\,\alpha^3\beta + 5\,\alpha^2\beta^2 + 3\,\alpha^2 + 2\,\alpha\beta + 2\,\alpha\beta^3 + 1 + 3\,\beta^4 + 2\,\beta^2}{2\,\alpha^2 + 2\,\alpha^4 + 2\,\beta^4 + \alpha^2\beta^2 + 2\,\alpha\beta + 1}$$

$$\neq \frac{f(\alpha,\beta)^2}{1 + \alpha^2 + \beta^2 + \alpha^2\beta^2 + \alpha^4 + \beta^4}$$

but, of course, their infima $= 1$ and their minimizers are the same!

Weighted norms $\|\vec{f}\|_{2,w}^2 = \sum_j w_j (\vec{f})_j^2$ for weights $w_i > 0$

$$\mathcal{N}_{2,w}^{[f]}(\alpha) = \frac{f(\alpha_1,\ldots,\alpha_n)^2}{\sum\limits_{(i_1,\ldots,i_n)\leq\deg(f)} \frac{1}{w_{i_1,\ldots,i_n}} \alpha_1^{2i_1}\cdots\alpha_n^{2i_n}}$$

$$\vec{\tilde{f}} = \vec{f} - \frac{\tau^T \vec{f}}{\tau^T \mathrm{Diag}(w)^{-1}\tau} \mathrm{Diag}(w)^{-1}\tau,$$

Weighted norms $\|\vec{f}\|_{2,w}^2 = \sum_j w_j (\vec{f})_j^2$ for weights $w_i > 0$

$$\mathcal{N}_{2,w}^{[f]}(\alpha) = \frac{f(\alpha_1,\ldots,\alpha_n)^2}{\sum_{(i_1,\ldots,i_n)\le\deg(f)} \frac{1}{w_{i_1,\ldots,i_n}} \alpha_1^{2i_1}\cdots\alpha_n^{2i_n}}$$

$$\tilde{\vec{f}} = \vec{f} - \frac{\tau^T \vec{f}}{\tau^T \mathrm{Diag}(w)^{-1}\tau} \mathrm{Diag}(w)^{-1}\tau,$$

$w_j \to \infty$: coefficient remains fixed, e.g., 0

Note: for $\alpha = 0$ cannot fix non-zero constant coefficient $\mathcal{N} = \frac{1}{0}$

$w_j \to 0$: coefficient is a don't care

$$\tilde{f}(x) = f(x) - \frac{f(\alpha)}{\alpha^i} x^i, \quad \alpha \neq 0$$

Digression: l^∞-norm [Hitz, Kaltofen, Lakshman 1999]

$$\mathcal{N}_\infty^{[f]}(\alpha) = \frac{|f(\alpha)|}{\sum_{i \le \deg(f)} |\alpha^i|}$$

Example $f(x) = x^2 + 1$

$$\inf_{\alpha \in \mathbb{R}} \mathcal{N}_\infty^{[x^2+1]}(\alpha) = \frac{2}{3} < 1, \quad \tilde{f} = \frac{1}{3}x^2 - \frac{2}{3}x + \frac{1}{3} = \frac{1}{3}(x-1)^2$$

Similar formulas for l^p-norms [Hitz 1999, Stetter 2000]

Example: 2 **simultaneous** equations

$f_1(x,y) = x^4 + y^4 + 1$ and $f_2(x,y) = x^2 + x^2y^2 - 2xy + 1$

Compute

$$\rho_2(f_1, f_2) = \inf_{\alpha,\beta\in\mathbb{R}} \frac{f_1(\alpha,\beta)^2 + f_2(\alpha,\beta)^2}{\sum_{0\le i+j\le 4} \alpha^{2i}\beta^{2j}}$$

Use $\partial_\alpha, \partial_\beta$ and Gröbner bases

cf. [Becker, Powers, and Wörmann 2000] $z^2F(x,y)+1=0$

and Mohab Safey El Din's RAGlib package

Example: 2 **simultaneous** equations

$f_1(x,y) = x^4 + y^4 + 1$ and $f_2(x,y) = x^2 + x^2y^2 - 2xy + 1$

Compute

$$\rho_2(f_1, f_2) = \inf_{\alpha,\beta\in\mathbb{R}} \frac{f_1(\alpha,\beta)^2 + f_2(\alpha,\beta)^2}{\sum_{0\le i+j\le 4} \alpha^{2i}\beta^{2j}}$$

Use $\partial_\alpha, \partial_\beta$ and Gröbner bases

cf. [Becker, Powers, and Wörmann 2000] $z^2F(x,y) + 1 = 0$

and Mohab Safey El Din's RAGlib package

Show Maple worksheet

Example: 2 **simultaneous** equations

$f_1(x,y) = x^4 + y^4 + 1$ and $f_2(x,y) = x^2 + x^2y^2 - 2xy + 1$

Compute

$$\rho_2(f_1, f_2) = \inf_{\alpha,\beta\in\mathbb{R}} \frac{f_1(\alpha,\beta)^2 + f_2(\alpha,\beta)^2}{\sum_{0\le i+j\le 4} \alpha^{2i}\beta^{2j}}$$

Use $\partial_\alpha, \partial_\beta$ and Gröbner bases

cf. [Becker, Powers, and Wörmann 2000] $z^2F(x,y) + 1 = 0$

and Mohab Safey El Din's RAGlib package

Exact SOS certificate for $\rho_2(f_1, f_2) > \frac{64597306998078108}{100000000000000000}$

Exact Certification of Optima

Problems with sum-of-squares certificates:

1. Numerical sum-of-squares yields "≥ 0" approximately!

2. Exact optimum is high-degree/large-height algebraic number

Exact Certification of Optima

Problems with sum-of-squares certificates:

1. Numerical sum-of-squares yields "≥ 0" approximately!

2. Exact optimum is high-degree/large-height algebraic number

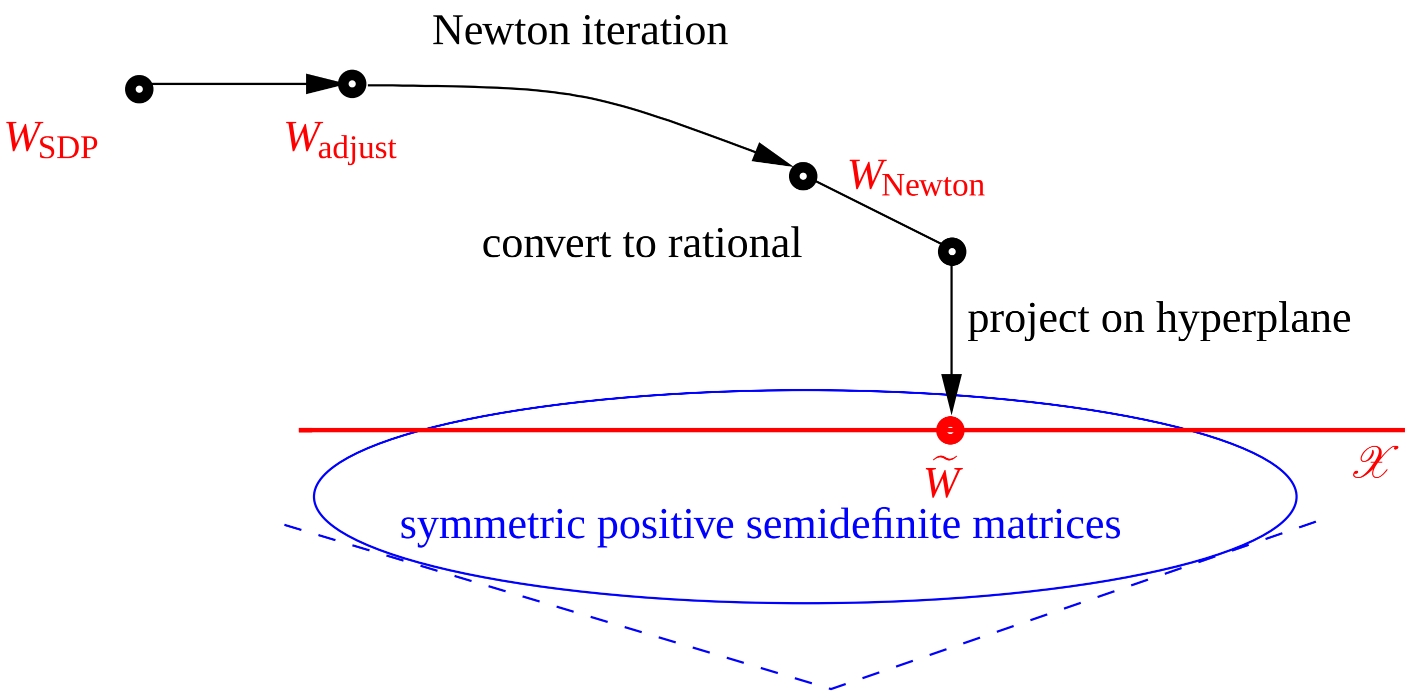
We certify a **rational** lower bound $\tilde{r} \lessapprox r \leq \dfrac{F}{G}$ via **rational** matrices $\widetilde{W}^{[1]}, \widetilde{W}^{[2]}$ so that the following conditions hold exactly:

$$(F(\mathbf{X}) - \tilde{r}G(\mathbf{X}))m_e(\mathbf{X})^T \cdot \widetilde{W}^{[2]} \cdot m_e(\mathbf{X}) = m_d(\mathbf{X})^T \cdot \widetilde{W}^{[1]} \cdot m_d(\mathbf{X}),$$

cf. [Nie et al. 2008]

$\widetilde{W}^{[1]} \succeq 0$, $\widetilde{W}^{[2]} \succeq 0$ $\quad (\dfrac{\text{SOS}}{\text{SOS}}$, 2 LMIs if $\tilde{r}$ is known)

Rationalizing a Sum-Of-Squares: "Easy Case"
Peyrl, Parrilo, '07, '08; Kaltofen, Li, Yang, Zhi, '08, '09



where the affine linear hyperplane is given by

$$\mathscr{X} = \{A \mid A^T = A, f(\mathbf{X}) = m_{\mathscr{G}}(\mathbf{X})^T \cdot A \cdot m_{\mathscr{G}}(\mathbf{X})\}$$

Example 1: Siegfried Rump's 2006 Model Problem

For $n = 1, 2, 3, \ldots$ compute the global minimum μ_n:

$$\mu_n = \min_{P,Q} \frac{\|PQ\|_2^2}{\|P\|_2^2 \|Q\|_2^2} \qquad \text{(rational function)}$$

$$\text{s. t. } P(Z) = \sum_{i=1}^{n} p_i Z^{i-1}, Q(Z) = \sum_{i=1}^{n} q_i Z^{i-1} \in \mathbb{R}[Z] \setminus \{0\}$$

Example 1: Siegfried Rump's 2006 Model Problem

For $n = 1, 2, 3, \ldots$ compute the global minimum μ_n:

$$\mu_n = \min_{P,Q} \frac{\|PQ\|_2^2}{\|P\|_2^2 \|Q\|_2^2} \qquad \text{(rational function)}$$

$$\text{s. t. } P(Z) = \sum_{i=1}^{n} p_i Z^{i-1}, Q(Z) = \sum_{i=1}^{n} q_i Z^{i-1} \in \mathbb{R}[Z] \setminus \{0\}$$

$$\Updownarrow$$

$$\mu_n = \min_{P,Q} \|PQ\|_2^2$$

$$\text{s. t. } \|P\|_2 = \|Q\|_2 = 1, \deg(P) \le n-1, \deg(Q) \le n-1$$

Local Minimum By Lagrangian Multipliers

Example 1: Siegfried Rump's 2006 Model Problem

For $n = 1, 2, 3, \ldots$ compute the global minimum μ_n :

$$\mu_n = \min_{P,Q} \frac{\|PQ\|_2^2}{\|P\|_2^2 \|Q\|_2^2} \qquad \text{(rational function)}$$

$$\text{s. t. } P(Z) = \sum_{i=1}^{n} p_i Z^{i-1}, Q(Z) = \sum_{i=1}^{n} q_i Z^{i-1} \in \mathbb{R}[Z] \setminus \{0\}$$

$$\Updownarrow$$

$$\frac{1}{\mu_n} = \max_{P,Q} B_{n-1}$$

$$\text{s. t. } \|P(Z)\|_2^2 \cdot \|Q(Z)\|_2^2 = B_{n-1} \|P(Z) \cdot Q(Z)\|_2^2$$

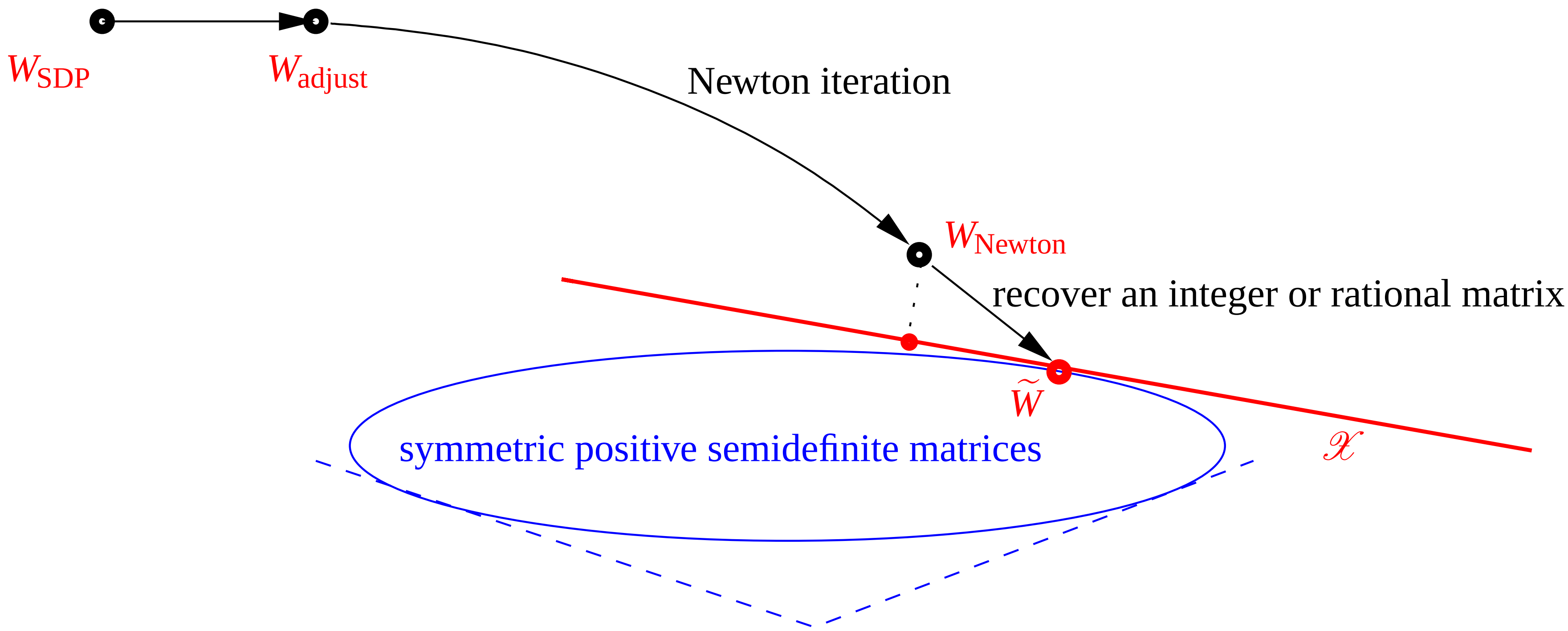
$$P, Q \in \mathbb{R}[Z] \setminus \{0\}, \deg(P) \leq n-1, \deg(Q) \leq n-1$$

Mignotte's factor coefficient bound: $\frac{1}{\mu_n} \leq \binom{2n-2}{n-1}^2$

Certified Lower Bounds by High Precision SDP [w. Feng Guo]

n	k	# iter	prec.	secs/iter	lower bound r_n	upper bound
4	2	50	4×15	0.71	0.0174291733214326528	0.0174291733214326529
5	1	50	4×15	2.03	0.0023395955481555912	0.0023395955481555913
6	2	50	4×15	1.76	0.00028973187527968192	0.0002897318752796819 3
7	1	75	5×15	11.36	0.0000341850698000828 4	0.0000341850698000828 5
8	2	75	5×15	12.49	0.0000039054356497557 2	0.0000039054356497557 3
9	1	75	5×15	84.12	0.436001653918104846 13e–06	0.436001653918104846 13e–06
10	2	75	5×15	92.79	0.478393956877097593 27e–07	0.478393956877097593 27e–07
11	1	85	5×15	622.03	0.517874909744699053 31e–08	0.517874909744699053 31e–08
12	2	85	5×15	634.48	0.5545881831163134761e–09	0.5545881831163134762e–09
13	1	100	5×15	3800.0	0.5886688081186609313 0e–10	0.5886688081186609313 0e–10
14	2	100	5×15	3800.00	0.6202444992053905021 9e–11	0.6202444992053905022 0e–11
15	1	120	6×15	15000.00	0.6494365418580951288 0e–12	0.6494365418580951288 0e–12
16	2	120	6×15	23000.00	0.6763604255822137905 7e–13	0.6763604255822137905 8e–13
17	1	70	6×15	72400.00	0.7011263189635532515 0e–14	0.7011263197014374158 5e–14
18	2	50	6×15	95720.00	0.7115460486506939698 8e–15	0.7238394479694387586 2e–15

Rationalizing a Sum-Of-Squares: “Hard Case”



where the affine linear hyperplane is tangent to the cone boundary

singular $\widetilde{W}$: real optimizers, fewer squares, missing terms

Proof of MCP conjecture for $n = 4$ (all rational coeff.'s)

6 polynomials of degree 8 with 8 variables:

2 polynomials are squares

1 polynomial is an SOS

3 polynomials are $\dfrac{\text{SOS}}{\text{weighted sum of squares of variables}}$

Examples From Literature

Example	The Denominator	#iter	prec.	#sq	secs
delzell	$2X_1^2 + 2X_2^2 + 2X_3^2$	*Null*	2×15	8	0.02
$motzkin(X_1, 3X_2, 3X_3)$	$4X_1^2 + 12X_2^2 + 21X_3^2$	19	1×15	5	0.304
$motzkin(X_1, 3X_2, 3X_3)$	$(X_1^2 + X_2^2 + X_3^2)^2$	96	10×15	7	17.217
leepstarr2	$15 + 20X_1^2 + 18X_2^2$	*Null*	1×15	9	0.344
laxlax	$X_1^2 + X_2^2 + X_3^2 + X_4^2$	*Null*	2×15	7	0.52
voronoi2	1	78	4×15	5	15.893

$$delzell(X_1, X_2, X_3, X_4) = X_1^4 X_2^2 X_4^2 + X_2^4 X_3^2 X_4^2 + X_1^2 X_3^4 X_4^2 - 3\,X_1^2 X_2^2 X_3^2 X_4^2 + X_3^8.$$

Concluding Observations

- Need help: prove that $f^2 - \rho_2(f)\|\tau\|_2^2$ is SOS.

Concluding Observations

- Need help: prove that $f^2 - \rho_2(f)\|\tau\|_2^2$ is SOS.

- Exact SOSes are an alternative to symbolic critical values computations in RAG

Concluding Observations

- Need help: prove that $f^2 - \rho_2(f)\|\tau\|_2^2$ is SOS.

- Exact SOSes are an alternative to symbolic critical values computations in RAG

- Win some—complexity on your side,
  tie some (MCP),
  lose some—rationality against you; [Reznick's challenges]

Concluding Observations

- Need help: prove that $f^2 - \rho_2(f)\|\tau\|_2^2$ is SOS.

- Exact SOSes are an alternative to symbolic critical values computations in RAG

- Win some—complexity on your side,
  tie some (MCP),
  lose some—rationality against you; [Reznick's challenges]

- Float inputs and exact output specs can be achieved

Thank you!