



## **Extending MOSEK with exponential cones**

ISMP Bordeaux 2018

[joachim.dahl@mosek.com](mailto:joachim.dahl@mosek.com)

[www.mosek.com](http://www.mosek.com)





Linear cone problem:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \in K, \end{array}$$

with  $K = K_1 \times K_2 \times \cdots \times K_p$  a product of proper cones.

Dual:

$$\begin{array}{ll} \text{maximize} & b^T y \\ \text{subject to} & c - A^T y = s \\ & s \in K^*, \end{array}$$

with  $K^* = K_1^* \times K_2^* \times \cdots \times K_p^*$ .





- *the nonnegative orthant*

$$K_l^n := \{x \in \mathbb{R}^n \mid x_j \geq 0, j = 1, \dots, n\},$$

- *the quadratic cone*

$$K_q^n = \{x \in \mathbb{R}^n \mid x_1 \geq (x_2^2 + \dots + x_n^2)^{1/2}\},$$

- *the rotated quadratic cone*

$$K_r^n = \{x \in \mathbb{R}^n \mid 2x_1x_2 \geq x_3^2 + \dots + x_n^2, x_1, x_2 \geq 0\}.$$

- *the semidefinite matrix cone*

$$K_s^n = \{x \in \mathbb{R}^{n(n+1)/2} \mid z^T \mathbf{mat}(x)z \geq 0, \forall z\}.$$





- *the three-dimensional power cone*

$$K_{\text{pow}}^{\alpha} = \{x \in \mathbb{R}^3 \mid x_1^{\alpha} x_2^{(1-\alpha)} \geq |x_3|, x_1, x_2 > 0\},$$

for  $0 < \alpha < 1$ .

- *the exponential cone*

$$K_{\text{exp}} = \text{cl}\{x \in \mathbb{R}^3 \mid x_1 \geq x_2 \exp(x_3/x_2), x_2 > 0\}.$$





Central path for homogenous model parametrized by  $\mu$ :

$$\begin{aligned}Ax_\mu - b\tau_\mu &= \mu(Ax - b\tau) \\s_\mu + A^T y_\mu - c\tau_\mu &= \mu(s + A^T y - c\tau) \\c^T x_\mu - b^T y_\mu + \kappa_\mu &= \mu(c^T x - b^T y + \kappa) \\s_\mu = -\mu F'(x_\mu), \quad x_\mu = -\mu F'_*(s_\mu), \quad \kappa_\mu \tau_\mu &= \mu,\end{aligned}$$

or equivalently

$$\begin{bmatrix} 0 & A & -b \\ -A^T & 0 & c \\ b^T & -c^T & 0 \end{bmatrix} \begin{bmatrix} y_\mu \\ x_\mu \\ \tau_\mu \end{bmatrix} - \begin{bmatrix} 0 \\ s_\mu \\ \kappa_\mu \end{bmatrix} = \mu \begin{bmatrix} r_p \\ r_d \\ r_g \end{bmatrix}$$
$$s_\mu = -\mu F'(x_\mu), \quad x_\mu = -\mu F'_*(s_\mu), \quad \kappa_\mu \tau_\mu = \mu,$$

$$r_p := Ax - b\tau, \quad r_d := c\tau - A^T y - s, \quad r_g := \kappa - c^T x + b^T y, \quad r_c := x^T s + \tau \kappa.$$



Following Tunçel [3] we consider a scaling  $W^T W \succ 0$ ,

$$v = Wx = W^{-T}s, \quad \tilde{v} = W\tilde{x} = W^{-T}\tilde{s}$$

where  $\tilde{x} := -F'_*(s)$  and  $\tilde{s} := -F'(x)$ . The centrality conditions

$$x = \mu\tilde{x}, \quad s = \mu\tilde{s}$$

can then be written symmetrically as

$$v = \mu\tilde{v},$$

and we linearize the centrality condition  $v = \mu\tilde{v}$  as

$$W\Delta x + W^{-T}\Delta s = -v + \mu\tilde{v}.$$





Let  $\mu := \frac{x^T s + \tau \kappa}{\nu + 1}$  with barrier parameter  $\nu$  and centering  $\gamma$ .

$$\begin{bmatrix} 0 & A & -b \\ -A^T & 0 & c \\ b^T & -c^T & 0 \end{bmatrix} \begin{bmatrix} \Delta y_c \\ \Delta x_c \\ \Delta \tau_c \end{bmatrix} - \begin{bmatrix} 0 \\ \Delta s_c \\ \Delta \kappa_c \end{bmatrix} = (\gamma - 1) \begin{bmatrix} r_p \\ r_d \\ r_g \end{bmatrix}$$

$$W \Delta x_c + W^{-T} \Delta s_c = \gamma \mu \tilde{v} - \nu, \quad \tau \Delta \kappa_c + \kappa \Delta \tau_c = \gamma \mu - \kappa \tau,$$

Constant decrease of residuals and complementarity:

$$\begin{aligned} Ax^+ - b\tau^+ &= \eta \cdot r_p, \\ c\tau^+ - A^T y^+ - s^+ &= \eta \cdot r_d, \\ b^T y^+ - c^T x^+ - \kappa^+ &= \eta \cdot r_g, \\ (x^+)^T s^+ + \tau^+ \kappa^+ &= \eta \cdot r_c, \end{aligned}$$

where  $z^+ := (z + \alpha \Delta z_c)$  and  $\eta = (1 - \alpha(1 - \gamma))$ .





Derivatives of  $s_\mu = -\mu F'(x_\mu)$ :

$$\dot{s}_\mu + \mu F''(x_\mu) \dot{x}_\mu = -F'(x_\mu),$$

$$\ddot{s}_\mu + \mu F''(x_\mu) \ddot{x}_\mu = -2F''(x_\mu) \dot{x}_\mu - \mu F'''(x_\mu) [\dot{x}_\mu, \dot{x}_\mu].$$

Since

$$\mu \dot{x}_\mu = -[F''(x_\mu)]^{-1} (F'(x_\mu) + \dot{s}_\mu) = x_\mu - [F''(x_\mu)]^{-1} \dot{s}_\mu,$$

we have

$$\mu F'''(x_\mu) [\dot{x}_\mu, \dot{x}_\mu] = \underbrace{F'''(x_\mu) [\dot{x}_\mu, x_\mu]}_{-2F''(x_\mu) \dot{x}_\mu} - F'''(x_\mu) [\dot{x}_\mu, (F''(x_\mu))^{-1} \dot{s}_\mu]$$

so

$$\ddot{s}_\mu + \mu F''(x_\mu) \ddot{x}_\mu = F'''(x_\mu) [\dot{x}_\mu, (F''(x_\mu))^{-1} \dot{s}_\mu].$$







Affine search-direction:

$$\begin{bmatrix} 0 & A & -b \\ -A^T & 0 & c \\ b^T & -c^T & 0 \end{bmatrix} \begin{bmatrix} \Delta y_a \\ \Delta x_a \\ \Delta \tau_a \end{bmatrix} - \begin{bmatrix} 0 \\ \Delta s_a \\ \Delta \kappa_a \end{bmatrix} = - \begin{bmatrix} r_p \\ r_d \\ r_g \end{bmatrix}$$

$$W\Delta x_a + W^{-T}\Delta s_a = -v, \quad \tau\Delta \kappa_a + \kappa\Delta \tau_a = -\kappa\tau,$$

satisfies

$$(\Delta x_a)^T \Delta s_a + \Delta \tau_a \Delta \kappa_a = 0.$$

Since

$$\dot{s}_\mu + \mu F''(x_\mu)\dot{x}_\mu = -F'(x_\mu) = s_\mu,$$

we interpret  $\Delta s_a = -\dot{s}_\mu$  and  $\Delta x_a = -\dot{x}_\mu$ .





From

$$\ddot{s}_\mu + \mu F''(x_\mu) \ddot{x}_\mu = F'''(x_\mu) [\dot{x}_\mu, (F''(x_\mu))^{-1} \dot{s}_\mu]$$

we define a corrector direction as

$$W \Delta x_{\text{cor}} + W^{-T} \Delta s_{\text{cor}} = \frac{1}{2} W^{-T} F'''(x) [\Delta x_a, (F''(x))^{-1} \Delta s_a].$$

Note that

$$\begin{aligned} s^T \Delta x_{\text{cor}} + x^T \Delta s_{\text{cor}} &= \frac{1}{2} x^T F'''(x) [\Delta x_a, (F''(x))^{-1} \Delta s_a] \\ &= -(\Delta x_a)^T \Delta s_a, \end{aligned}$$

condition for constant decrease of complementarity.





- Linear case,

$$\frac{1}{2}F'''(x)[\Delta x_a, (F''(x))^{-1}\Delta s_a] = -\mathbf{diag}(x)^{-1}\mathbf{diag}(\Delta x_a)\Delta s_a,$$

- Semidefinite case,

$$\begin{aligned}\frac{1}{2}F'''(x)[\Delta x_a, (F''(x))^{-1}\Delta s_a] &= -\frac{1}{2}x^{-1}\Delta x_a\Delta s_a - \frac{1}{2}\Delta s_a\Delta x_a x^{-1} \\ &= -(x^{-1}) \circ (\Delta x_a\Delta s_a),\end{aligned}$$

- Second-order cone case,

$$F'''(x)[(F''(x))^{-1}u] = -\frac{2}{x^T Q x}(ux^T Q + Qxu^T - (x^T u)Q)$$

for  $Q = \mathbf{diag}(1, -1, \dots, -1)$ . Then

$$F'''(x)[(F''(x))^{-1}u]e = -2(x^{-1} \circ u).$$





A combined centering-corrector direction:

$$\begin{bmatrix} 0 & A & -b \\ -A^T & 0 & c \\ b^T & -c^T & 0 \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta x \\ \Delta \tau \end{bmatrix} - \begin{bmatrix} 0 \\ \Delta s \\ \Delta \kappa \end{bmatrix} = (\gamma - 1) \begin{bmatrix} r_p \\ r_d \\ r_g \end{bmatrix}$$

$$W\Delta x + W^{-1}\Delta s = \gamma\mu\tilde{v} - v + \frac{1}{2}W^{-T}F'''(x)[\Delta x_a, (F''(x))^{-1}\Delta s_a],$$
$$\tau\Delta\kappa + \kappa\Delta\tau = \gamma\mu - \tau\kappa - \Delta\tau_a\Delta\kappa_a.$$

All residuals and complementarity decrease by  $\eta$ .





## Theorem (Schnabel [2])

Let  $S, Y \in \mathbb{R}^{n \times p}$  have full rank  $p$ . Then there exists  $H \succ 0$  such that  $HS = Y$  if and only if  $Y^T S \succ 0$ .

As a consequence

$$H = Y(Y^T S)^{-1} Y^T + ZZ^T$$

where  $S^T Z = 0$ ,  $\mathbf{rank}(Z) = n - p$ . We have  $n = 3$ ,  $p = 2$  and

$$S := (x \quad \tilde{x}), \quad Y := (s \quad \tilde{s}),$$

with

$$\det(Y^T S) = \nu^2(\mu\tilde{\mu} - 1) \geq 0$$

vanishing only on the central path.





Any scaling with  $n = 3$  satisfies

$$W^T W = Y(Y^T S)^{-1} Y^T + z z^T$$

where  $(x \quad \tilde{x})^T z = 0$ ,  $z \neq 0$ . Expanding the BFGS update [2]

$$H^+ = H + Y(Y^T S)^{-1} Y^T - H S (S^T H S)^{-1} S^T H,$$

for  $H \succ 0$  gives the scaling by Tunçel [3] and Myklebust [1], *i.e.*,

$$z z^T = H - H S (S^T H S)^{-1} S^T H,$$

with  $H = \mu F''(x)$ .





Nesterov's long-step Hessian estimation property holds if

$$F'''(x)[u] \preceq 0, \quad \forall x \in \text{int}(K), \forall u \in K.$$

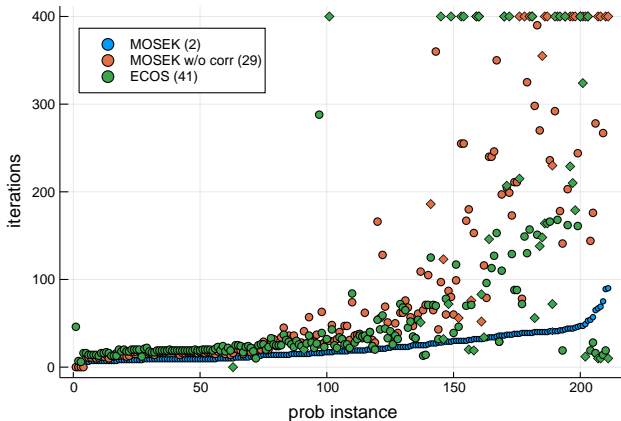
We have

$$F'''([1; 1; -1])[u] = u_1 \begin{bmatrix} -9 & 6 & 3 \\ 6 & -5 & -3 \\ 3 & -3 & -2 \end{bmatrix} + u_2 \begin{bmatrix} 6 & -5 & -3 \\ -5 & 2 & 3 \\ -3 & 3 & 2 \end{bmatrix} + u_3 \begin{bmatrix} 3 & -3 & -2 \\ -3 & -3 & 2 \\ -2 & 2 & 2 \end{bmatrix}.$$

Not negative semidefinite for all  $u \in K$ .



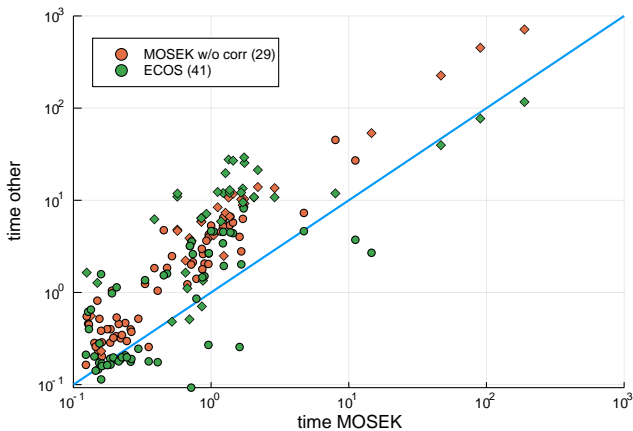
# Comparing MOSEK and ECOS conic solvers



Iteration counts for different exponential cone problems. Failures marked with  $\diamond$ .

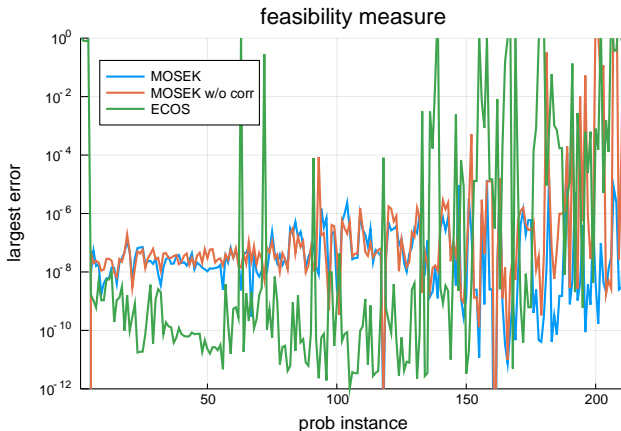


# Comparing MOSEK and ECOS conic solvers

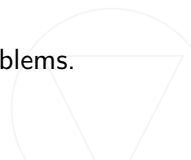


Solution time for different exponential cone problems. Failures marked with  $\diamond$ .

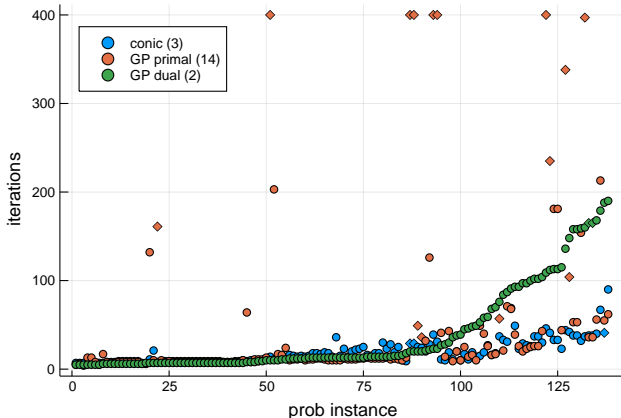
# Comparing MOSEK and ECOS conic solvers



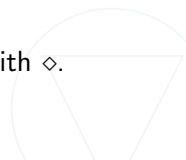
Feasibility measures for different exponential cone problems.



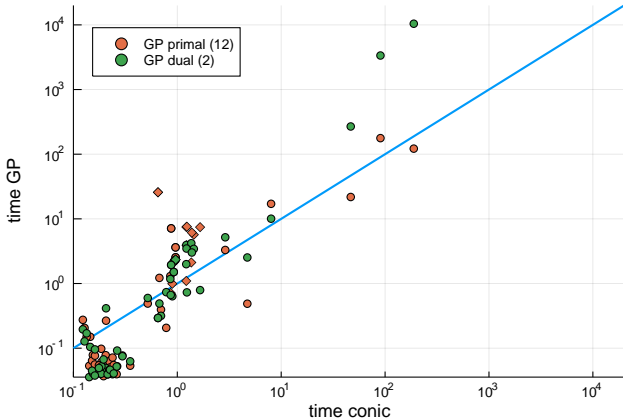
# Comparing MOSEK conic and MOSEK GP solvers



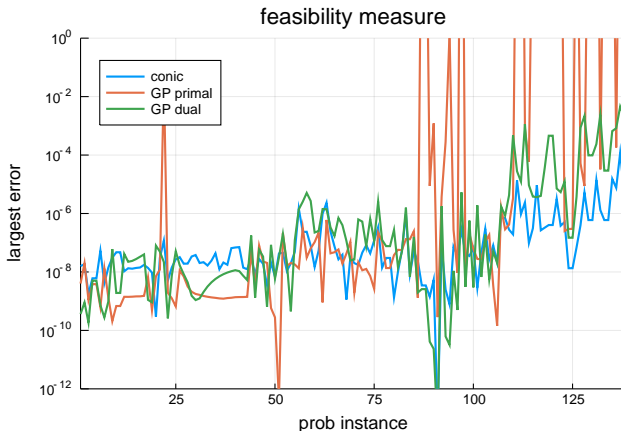
Iteration counts for different GPs. Failures marked with  $\diamond$ .



# Comparing MOSEK conic and MOSEK GP solvers



# Comparing MOSEK conic and MOSEK GP solvers



Feasibility measures for different GPs.





- [1] T. Myklebust and L. Tunçel.  
Interior-point algorithms for convex optimization based on primal-dual metrics.  
Technical report, 2014.
- [2] R. B. Schnabel.  
Quasi-newton methods using multiple secant equations.  
Technical report, Colorado Univ., Boulder, Dept. Comp. Sci., 1983.
- [3] L. Tunçel.  
Generalization of primal-dual interior-point methods to convex optimization problems in conic form.  
*Foundations of Computational Mathematics*, 1:229–254, 2001.

