

On the Conditioning of Polynomial Eigenproblem Linearizations

Françoise Tisseur
School of Mathematics
University of Manchester

`ftisseur@ma.man.ac.uk`

`http://www.ma.man.ac.uk/~ftisseur/`

Joint work with **Nicholas J. Higham**
and **D. Steven Mackey**.

Polynomial Eigenproblem

$$P(\lambda) = \sum_{i=0}^m \lambda^i A_i, \quad A_i \in \mathbb{C}^{n \times n}, \quad A_m \neq 0.$$

P assumed **regular** ($\det P(\lambda) \not\equiv 0$).

Find scalars λ and nonzero vectors x and y satisfying $P(\lambda)x = 0$ and $y^*P(\lambda) = 0$.

- Standard eigenvalue problem: $Ax = \lambda x$.
- Generalized eigenvalue problem (GEP): $Ax = \lambda Bx$.
- Quadratic eigenvalue problem (QEP):
 $(\lambda^2 M + \lambda C + K)x = 0$.

Linearizations

The pencil

$$L(\lambda) = \lambda X + Y, \quad X, Y \in \mathbb{C}^{mn \times mn}$$

is a **linearization** of $P(\lambda) = \sum_{i=0}^m \lambda^i A_i$ if

$$E(\lambda)L(\lambda)F(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{(m-1)n} \end{bmatrix}$$

for some **unimodular** $E(\lambda)$ and $F(\lambda)$.

Standard way of solving $P(\lambda)x = 0$:

- ▶ Linearize $P(\lambda)$ into $L(\lambda) = \lambda X + Y$.
- ▶ Solve generalized eigenproblem $L(\lambda)z = 0$.

\mathbb{L}_1 and \mathbb{L}_2 Linearizations

$$A := [\lambda^{m-1}, \lambda^{m-2}, \dots, 1]^T.$$

Mackey, Mackey, Mehl & Mehrmann (2005) define

$$\begin{aligned}\mathbb{L}_1(P) &= \{ L(\lambda) : L(\lambda)(A \otimes I_n) = v \otimes P(\lambda), v \in \mathbb{C}^m \}, \\ \mathbb{L}_2(P) &= \{ L(\lambda) : (A^T \otimes I_n)L(\lambda) = w^T \otimes P(\lambda), w \in \mathbb{C}^m \}.\end{aligned}$$

They show that

- \mathbb{L}_1 and \mathbb{L}_2 are vector spaces.
- Almost all pencils in \mathbb{L}_1 and \mathbb{L}_2 are linearizations.

Quadratic case ($m = 2$): $L = \lambda X + Y \in \mathbb{L}_1(P)$ iff

$$\begin{bmatrix} v_1 M & v_1 C & v_1 K \\ v_2 M & v_2 C & v_2 K \end{bmatrix} = \begin{bmatrix} X_{11} & X_{12} + Y_{11} & Y_{12} \\ X_{21} & X_{22} + Y_{21} & Y_{22} \end{bmatrix}.$$

\mathbb{L}_1 and \mathbb{L}_2 Linearizations cont.

Recall

$$\mathbb{L}_1(P) = \{ L(\lambda) : L(\lambda)(A \otimes I_n) = v \otimes P(\lambda), v \in \mathbb{C}^m \}.$$

If x is an eigenvector of P then

$$L(\lambda)(A \otimes x) = L(\lambda)(A \otimes I_n)(1 \otimes x) = (v \otimes P(\lambda))(1 \otimes x) = v \otimes P(\lambda)x.$$

So (x, λ) is an ei'pair of P iff $(A \otimes x, \lambda)$ is an ei'pair of L .

- **Right** eigenvectors of P can be recovered from **right** eigenvectors of linearizations in \mathbb{L}_1 .
- **Left** eigenvectors of P can be recovered from **left** eigenvectors of linearizations in \mathbb{L}_2 .

$\mathbb{DL}(P)$ Linearizations

Mackey, Mackey, Mehl & Mehrmann (2005) define

$$\mathbb{DL}(P) = \mathbb{L}_1(P) \cap \mathbb{L}_2(P).$$

They show that

- $L \in \mathbb{DL}(P)$ iff $w = v$ in the definitions of \mathbb{L}_1 and \mathbb{L}_2 .
- $\mathbb{DL}(P)$ is a vector space of dimension m .

In this work we focus on linearizations in $\mathbb{DL}(P)$.

Example: For $Q(\lambda) = \lambda^2 A + \lambda B + C$, $\mathbb{DL}(Q)$ is the pencils

$$L(\lambda) = \lambda \begin{bmatrix} v_1 A & v_2 A \\ v_2 A & v_2 B - v_1 C \end{bmatrix} + \begin{bmatrix} v_1 B - v_2 A & v_1 C \\ v_1 C & v_2 C \end{bmatrix}, \quad v \in \mathbb{C}^2.$$

Key Questions

- ★ For given λ which $L \in \mathbb{DL}(P)$ has the **smallest condition number** for λ (minimization over v).
- ★ How does this minimal condition number compare with **condition number of P** for λ ?
- ★ How well conditioned is L for $v = e_1$ and $v = e_m$?

Condition Number for $P(\lambda)$

Simple λ , $P(\lambda)x = 0$, $y^*P(\lambda) = 0$.

$$\kappa_P(\lambda) = \limsup_{\epsilon \rightarrow 0} \left\{ \frac{|\Delta\lambda|}{\epsilon|\lambda|} : (P(\lambda + \Delta\lambda) + \Delta P(\lambda + \Delta\lambda))(x + \Delta x) = 0, \right. \\ \left. \|\Delta A_i\|_2 \leq \epsilon \omega_i, i = 0:m \right\}.$$

F.T. (2000) showed

$$\kappa_P(\lambda) = \frac{(\sum_{i=0}^m |\lambda|^i \omega_i) \|y\|_2 \|x\|_2}{|\lambda| |y^* P'(\lambda) x|}.$$

Problems

- κ_P not defined for $\lambda = 0$.
- κ_P not defined for $\lambda = \infty$.

Homogeneous Form

$$P(\alpha, \beta) = \sum_{i=0}^m \alpha^i \beta^{m-i} A_i \quad (\lambda \equiv \alpha/\beta).$$

E'vals are pairs $(\alpha, \beta) \neq (0, 0)$ s.t. $\det P(\alpha, \beta) = 0$.

Dedieu (1997) and Dedieu & Tisseur (2003) define **condition number of (α, β)** as norm of a **condition operator** $K(\alpha, \beta)$ that is the differential of map from the $(m + 1)$ -tuple (A_0, \dots, A_m) to (α, β) in projective space:

$$\kappa_P(\alpha, \beta) = \max_{\|\Delta A\| \leq 1} \frac{\|K(\alpha, \beta) \Delta A\|_2}{\|[\alpha, \beta]\|_2},$$

where $\Delta A \equiv (\Delta A_0, \dots, \Delta A_m)$.

Condition Number for Homogeneous Form

Can show that, with suitable normalizations,

$$|\theta((\alpha, \beta), (\tilde{\alpha}, \tilde{\beta}))| \leq \kappa_P(\alpha, \beta) \|\Delta A\| + o(\|\Delta A\|).$$

Take sine of both sides to get bound for **chordal distance** between (α, β) and $(\tilde{\alpha}, \tilde{\beta})$ (Stewart & Sun, 1990).

Theorem 1 With $\|A\| := \|[\omega_0^{-1} A_0, \dots, \omega_m^{-1} A_m]\|_F$,

$$\kappa_P(\alpha, \beta) = \left(\sum_{i=0}^m |\alpha|^{2i} |\beta|^{2(m-i)} \omega_i^2 \right)^{1/2} \frac{\|x\|_2 \|y\|_2}{|y^*(\bar{\beta} D_\alpha P - \bar{\alpha} D_\beta P)|_{(\alpha, \beta) x}},$$

where $D_\alpha \equiv \frac{\partial}{\partial \alpha}$, $D_\beta \equiv \frac{\partial}{\partial \beta}$.

Condition Number κ_L (1)

$$A_{\alpha,\beta} = [\alpha^{m-1}, \alpha^{m-2}\beta, \dots, \beta^{m-1}]^T.$$

$$L(\alpha, \beta)z = 0, \quad w^*L(\alpha, \beta) = 0,$$

$$w = \bar{A}_{\alpha,\beta} \otimes y, \quad z = A_{\alpha,\beta} \otimes x.$$

$$\kappa_L(\alpha, \beta) = \sqrt{|\alpha|^2\omega_X^2 + |\beta|^2\omega_Y^2} \frac{\|w\|_2 \|z\|_2}{|w^*(\bar{\beta}D_\alpha L - \bar{\alpha}D_\beta L)|_{(\alpha,\beta)} z|}$$

$$L \in \mathbb{L}_1 \Rightarrow L(\alpha, \beta)(A_{\alpha,\beta} \otimes I_n) = v \otimes P(\alpha, \beta).$$

Differentiate wrt α :

$$D_\alpha L(\alpha, \beta)(A_{\alpha,\beta} \otimes I_n) + L(\alpha, \beta)(D_\alpha A_{\alpha,\beta} \otimes I_n) = v \otimes D_\alpha P(\alpha, \beta).$$

Multiplying on left by w^* and on right by $1 \otimes x$:

$$\begin{aligned} w^*D_\alpha L(\alpha, \beta)z &= A_{\alpha,\beta}^T v \otimes y^* D_\alpha P(\alpha, \beta)x \\ &= A_{\alpha,\beta}^T v \cdot y^* D_\alpha P(\alpha, \beta)x. \end{aligned}$$

Condition Number κ_L (2)

Theorem 2 For any pencil $L(\alpha, \beta) = \alpha X + \beta Y \in \text{DL}(P)$,

$$\kappa_L(\alpha, \beta) = \frac{\sqrt{|\alpha|^2 \omega_X^2 + |\beta|^2 \omega_Y^2} \|A_{\alpha, \beta}\|_2 \|x\|_2 \|y\|_2}{|p(\alpha, \beta; v)| |y^*(\bar{\beta} D_\alpha P - \bar{\alpha} D_\beta P)|_{(\alpha, \beta) x}},$$

where $p(\alpha, \beta; v) = A_{\alpha, \beta}^T v = \sum_{i=1}^m v_i \alpha^{m-i} \beta^{i-1}$.

- **Second term** in denominator is same as in $\kappa_P(\alpha, \beta)$ expression.
 - $\kappa_L(\alpha, \beta)$ is finite iff (α, β) is not a zero of $p(\alpha, \beta; v)$.
- Cf. M^4 paper:** $L(\lambda)$ is a linearization for $P(\lambda)$ iff no eigenvalue of P is a root of $p(\lambda; v)$.

Minimizing the Condition Number κ_L (1)

Pencil $L(\lambda) \in \text{IDIL}(P)$ defined by vector $v \in \mathbb{C}^m$.

Aim: minimize $\kappa_L(\alpha, \beta)$ over all v (with $\|v\|_2 = 1$ wlog).

Take

$$\omega_X = \|X\|_2, \quad \omega_Y = \|Y\|_2.$$

Can bound **numerator** of κ_L above and below.

To maximize **denominator** of κ_L , **maximize $p(\alpha, \beta; v)$** .

Cauchy–Schwarz gives

$$v_* = \frac{\overline{\Lambda}_{\alpha, \beta}}{\|\Lambda_{\alpha, \beta}\|_2}.$$

Important special cases:

$$\begin{aligned} (\alpha, \beta) = (1, 0), \quad \lambda = \infty &\Rightarrow v_* = e_1, \\ (\alpha, \beta) = (0, 1), \quad \lambda = 0 &\Rightarrow v_* = e_m. \end{aligned}$$

Minimizing the Condition Number κ_L (2)

Theorem 3 Consider pencils $L = \alpha X + \beta Y \in \text{DL}(P)$. Set weights $\omega_X = \|X\|_2$, $\omega_Y = \|Y\|_2$. Then

$$\kappa_L(\alpha, \beta; e_1) \leq \rho m^{3/2} \inf_v \kappa_L(\alpha, \beta, v) \text{ if } A_0 \text{ nonsingular, } |\alpha| \geq |\beta|,$$

$$\kappa_L(\alpha, \beta; e_m) \leq \rho m^{3/2} \inf_v \kappa_L(\alpha, \beta, v) \text{ if } A_m \text{ nonsingular, } |\alpha| \leq |\beta|,$$

where

$$\rho = \frac{\max_i \|A_i\|_2}{\min(\|A_0\|_2, \|A_m\|_2)}.$$

- For $\rho = O(1)$, **one of $v = e_1$ and $v = e_m$ gives near optimal κ_L** for λ .
- **Wrong choice** of $v = e_1$ or $v = e_m$ can be **disastrous**:
 $\kappa_L(0, \beta, e_1) = \infty$, $\kappa_L(\alpha, 0, e_m) = \infty$.

QEP Case

For $P(\lambda) = \lambda^2 A + \lambda B + C$,

$$\begin{aligned} v = e_1 &\Rightarrow L_1(\lambda) = \lambda \begin{bmatrix} A & 0 \\ 0 & -C \end{bmatrix} + \begin{bmatrix} B & C \\ C & 0 \end{bmatrix}, \\ v = e_m &\Rightarrow L_2(\lambda) = \lambda \begin{bmatrix} 0 & A \\ A & B \end{bmatrix} + \begin{bmatrix} -A & 0 \\ 0 & C \end{bmatrix}. \end{aligned}$$

F.T. (2000) showed that $\|A\|_2 = \|B\|_2 = \|C\|_2 = 1 \Rightarrow$

$$\kappa_{L_1}(\lambda) \leq \kappa_{L_2}(\lambda) \text{ for } |\lambda| \geq 2^{1/2},$$

$$\kappa_{L_1}(\lambda) \geq \kappa_{L_2}(\lambda) \text{ for } |\lambda| \leq 2^{-1/2}.$$

Our analysis shows L_1 and L_2 are **near-optimal over DL(P) for arbitrary m** for all eigenvalues with $|\lambda| \geq 1$ and $|\lambda| \leq 1$, resp., provided $\rho = O(1)$.

Optimal κ_L Versus κ_P

Theorem 4 Let (α, β) be a simple eigenvalue of P . Then

$$\frac{1}{\rho} \leq \frac{\inf_v \kappa_L(\alpha, \beta; v)}{\kappa_P(\alpha, \beta)} \leq m^2 \rho.$$

If $\rho = O(1)$:

- Best conditioned $L \in \mathbb{DL}(P)$ for a given λ is about as well conditioned as P itself for λ .

Despite perturbations to L not respecting the block structure of X and Y !

- Combined with Thm. 6:
one of pencils with $v = e_1$ and $v = e_m$ is about as well conditioned as P itself for λ .

Quadratic Case: $Q(\lambda) = \lambda^2 A + \lambda B + C$

Write $a = \|A\|_2$, $b = \|B\|_2$, $c = \|C\|_2$.

For optimality of $\kappa_L(\alpha, \beta)$ for $v = e_1$, $v = e_2$ need $\rho = O(1)$,
i.e.

$$b \lesssim \max(a, c) \quad \text{and} \quad a \approx c.$$

With the scaling : $\lambda^2 A + \lambda B + C \rightarrow \mu^2(\gamma^2 A) + \mu(\gamma B) + C$
($\lambda = \mu\gamma$), $\gamma = \sqrt{c/a}$ [Fan, Lin & Van Dooren, 2004],
it suffices that

$$b \lesssim \sqrt{ac}. \quad (*)$$

(*) is true for elliptic QEPs.

Companion Linearizations (1)

$C_i(\lambda) = \lambda X + Y_i$ with $X = \text{diag}(A_m, I_n, \dots, I_n)$,

$$Y_1 = \begin{bmatrix} A_{m-1} & A_{m-2} & \dots & A_0 \\ -I_n & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & -I_n & 0 \end{bmatrix}, \quad Y_2 = \begin{bmatrix} A_{m-1} & -I_n & \dots & 0 \\ A_{m-2} & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & -I_n \\ A_0 & 0 & \dots & 0 \end{bmatrix}.$$

- $C_1 \in \mathbb{L}_1(P)$ ($v = e_1$) and $C_2 \in \mathbb{L}_2(P)$ ($w = e_1$), but $C_1, C_2 \notin \mathbb{DL}(P)$.
- C_1 and C_2 **always** linearizations.
- Know left e'vec of C_1 in terms of left e'vec of P .
- Have formula for $\kappa_{C_1}(\lambda)$.

Companion Linearizations (2)

Find that κ_{C_1}/κ_P and $\kappa_{C_1}/\kappa_L(\lambda; v_*)$ depends on

- ratios $\|w\|_2/\|y\|_2$ of norms of left ei'vecs. of C_1 and P ,
- rational functions of the $\|A_i\|_2$'s and λ .

Conclude that

- If $\|A_i\| \approx 1, i = 0:m$ then $\kappa_{C_1} \approx \kappa_P$ and $\kappa_{C_1} \approx \kappa_L(\lambda; v_*)$.
- If $\|w\|_2/\|y\|_2 \gg 1$ or if $\|A_i\|_2 \gg 1, i = 0:m$, then $\kappa_{C_1} \gg \kappa_L(\lambda; v_*)$ is possible.

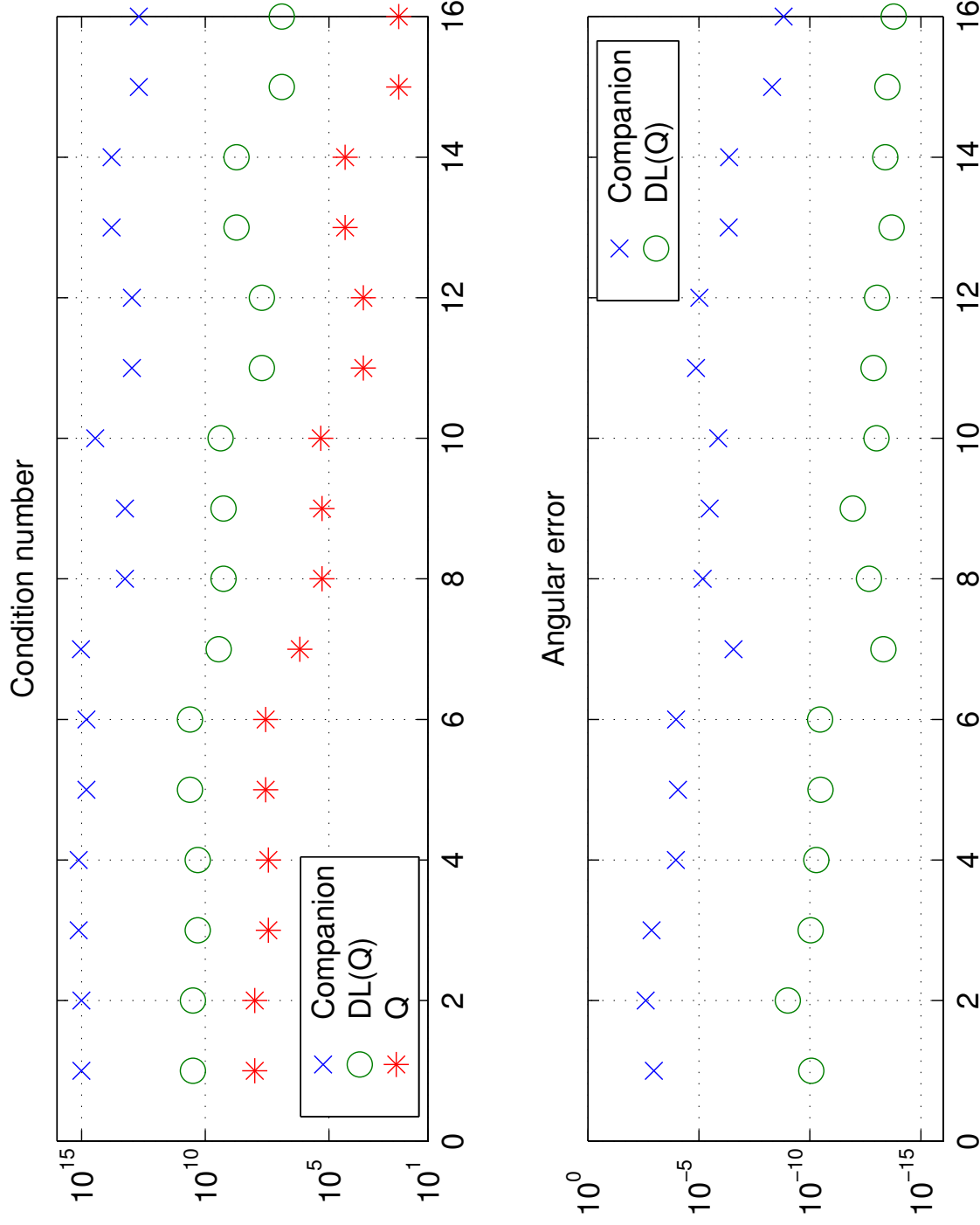
Same comments apply to C_2 since

$$\kappa_{C_2(P)}(\lambda) = \kappa_{C_1(P^T)}(\lambda).$$

Example 1: Nuclear Power Plant (1)

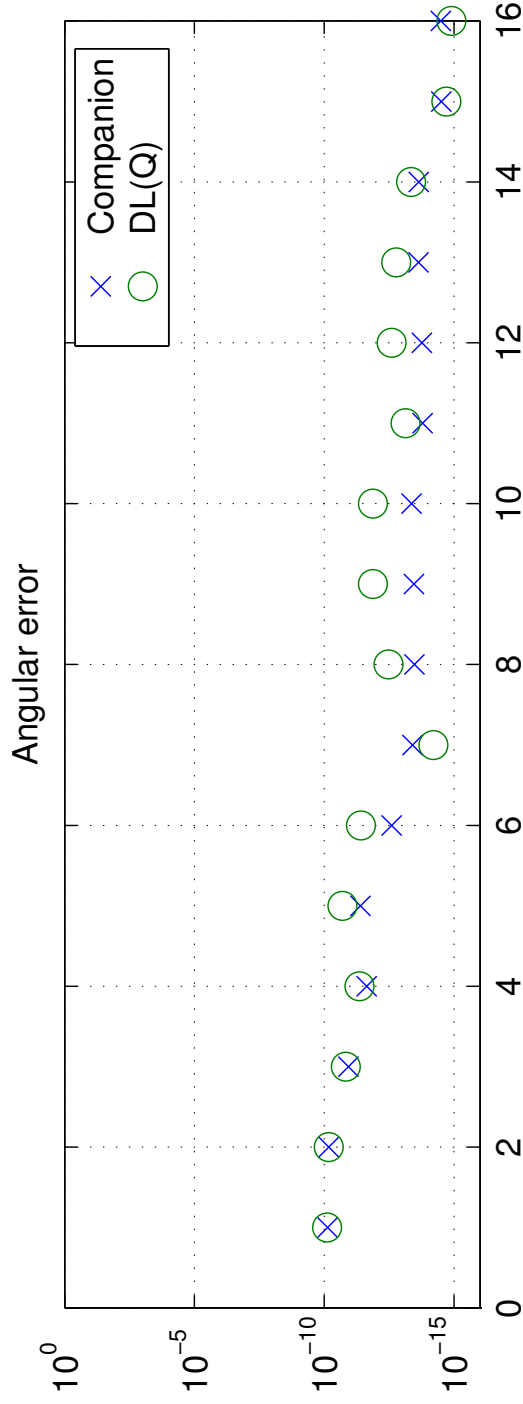
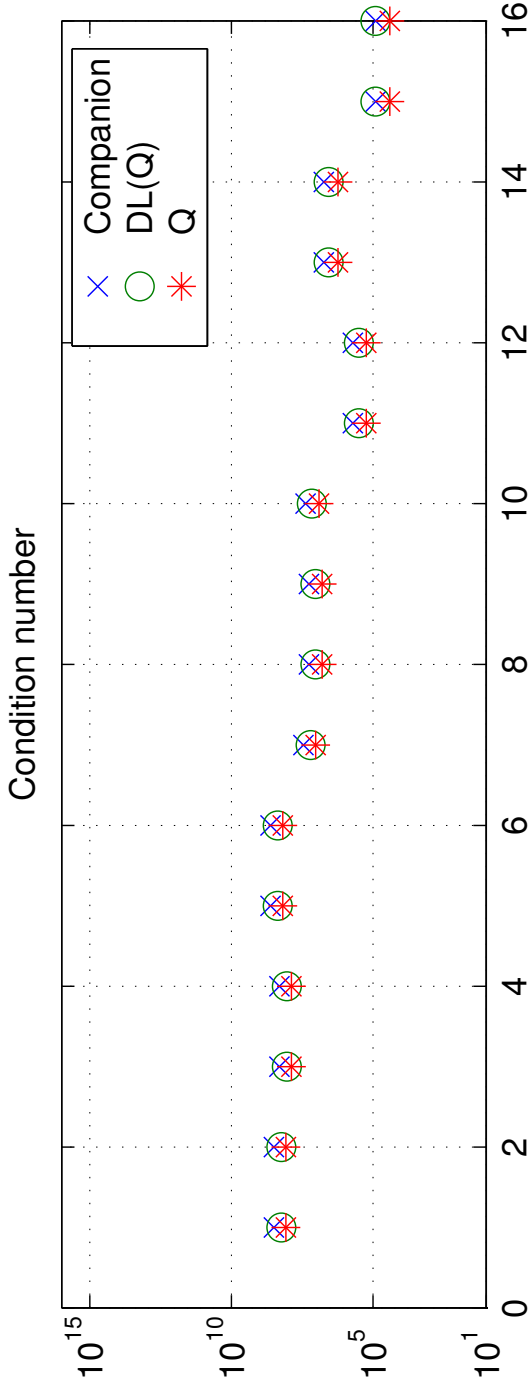
$n = 8$; $\|A\|_2 = 2.3 \times 10^8$, $\|B\|_2 = 4.3 \times 10^{10}$, $\|C\|_2 = 1.7 \times 10^{13}$.

$|\lambda| \in (17, 362)$; **Unscaled**; $v = e_1$, $\rho = 7 \times 10^4$.



Example 1: Nuclear Power Plant (2)

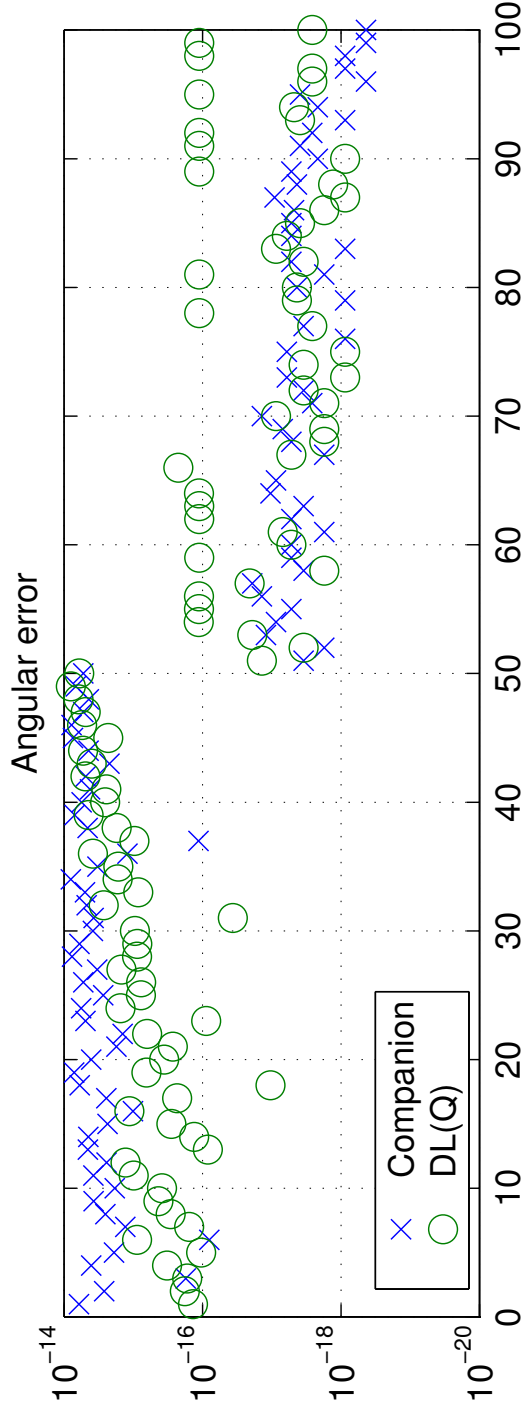
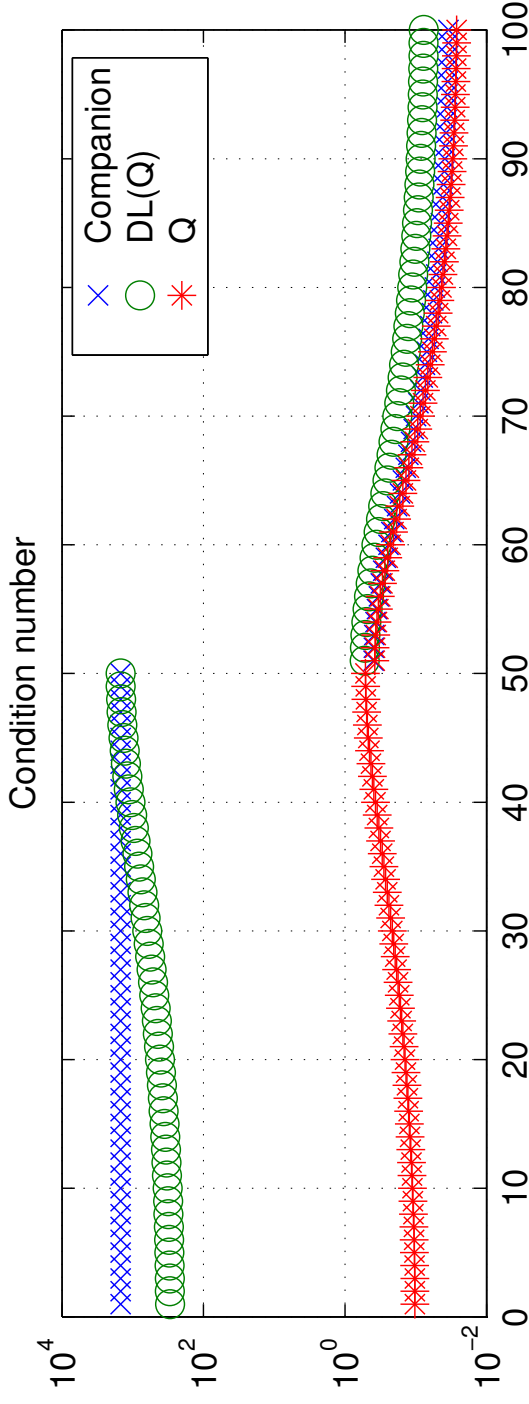
Scaled; $v = e_2$, $\rho = 1$.



Example 2: Damped Mass-Spring (1)

$n = 50$; $50 \ |\lambda| \in (-320, -6.4)$, $50 \ \lambda \approx -1.5 \times 10^{-2}$.

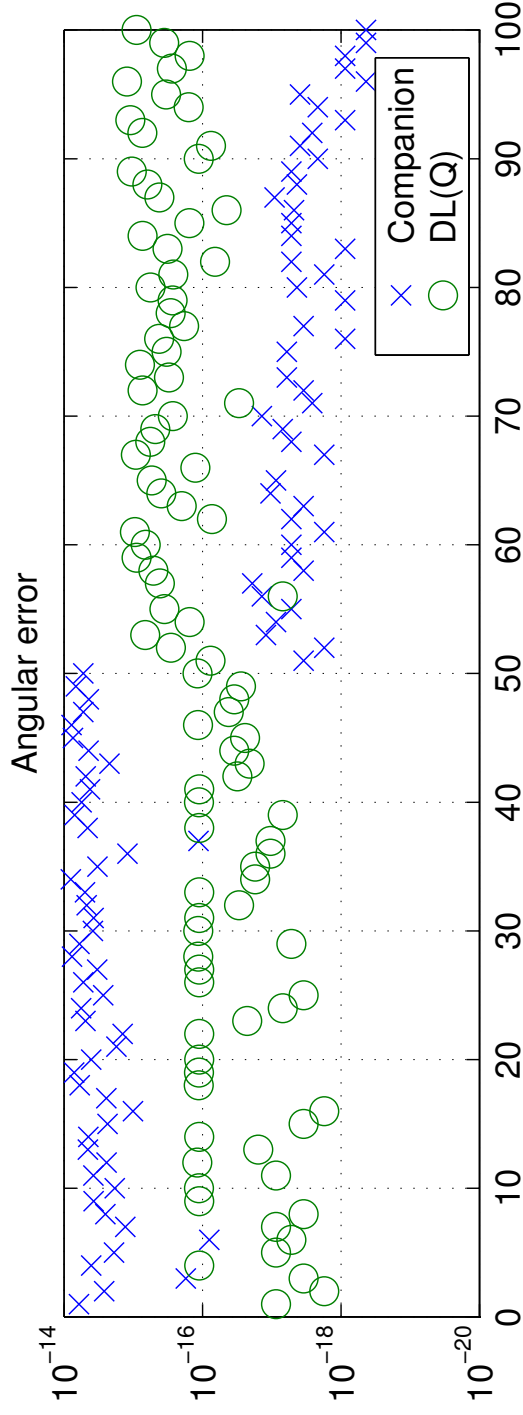
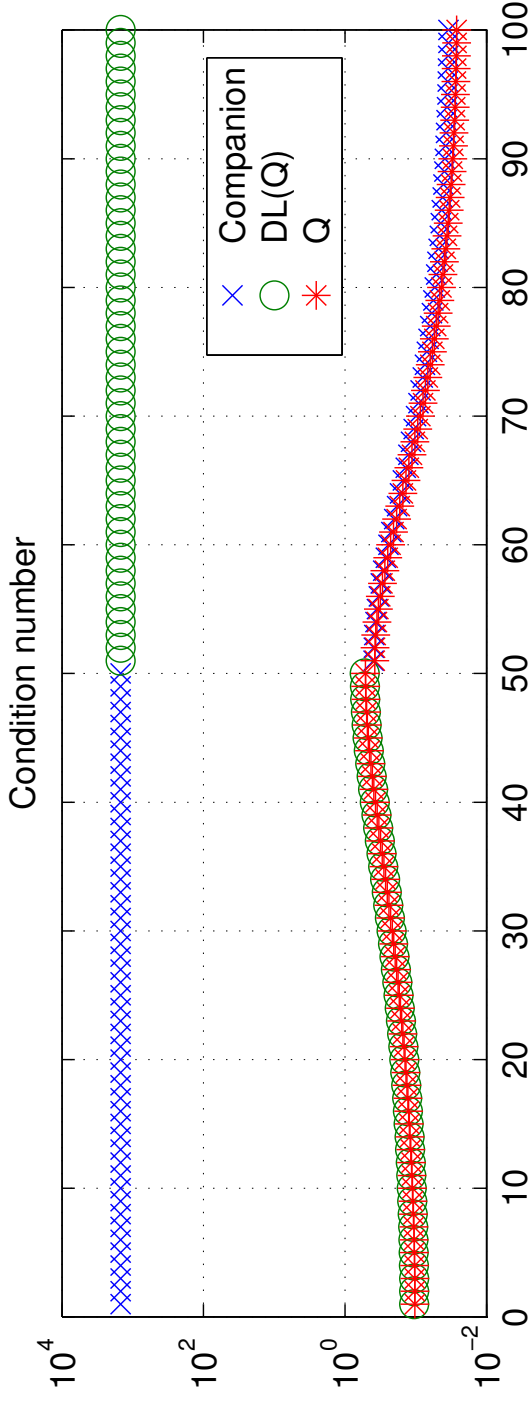
Unscaled; $v = e_1$, $\rho = 320$.



Example 2: Damped Mass-Spring (2)

$n = 50$; $50 \ |\lambda| \in (-320, -6.4)$, $50 \ \lambda \approx -1.5 \times 10^{-2}$.

Unscaled; $v = e_2$, $\rho = 320$.



Conclusions

$$P(\lambda) = \sum_{i=0}^m \lambda^i A_i, \quad L(\lambda) = \lambda X + Y \in \mathbb{DL}(P).$$

Analyzed ei'val conditioning of linearizations of $\mathbb{DL}(P)$.

- ★ Showed that if $\max_i \|A_i\|_2 \approx \min(\|A_0\|_2, \|A_m\|_2)$ then L for $v = e_1$ and $v = e_m$ **optimally conditioned** for $|\lambda| \geq 1$ and $|\lambda| \leq 1$, resp. **and** as well conditioned as P .
- ★ For quadratics, assumption on coeffs weakened to $\|B\|_2 \lesssim \sqrt{\|A\|_2 \|C\|_2}$ by use of scaling.
- ★ Companion linearizations can be poorly conditioned.

Justification for solving $P(\lambda)x = 0$ by linearization.