Department of Electrical Engineering (ESAT)
Faculty of Engineering
Katholieke Universiteit Leuven

Convexification of system identification methods

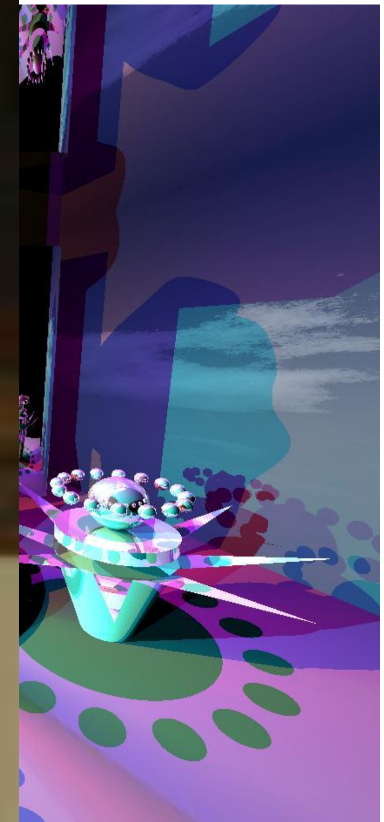
You don’t need eyes to see
You need vision

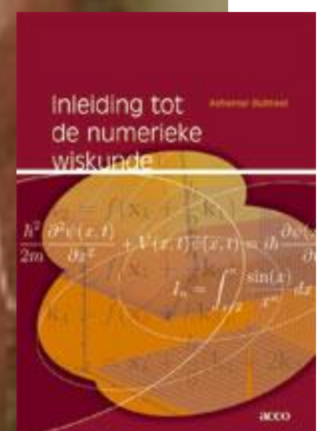
Bart De Moor
ESAT-SCD
Katholieke Universiteit Leuven

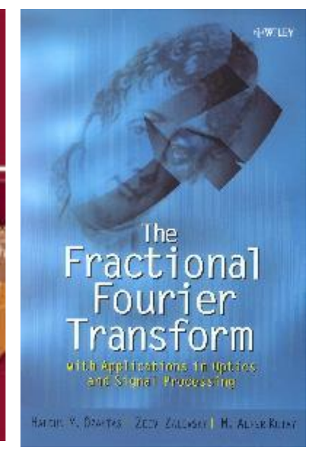
A: Kasteelpark Arenberg 10, B-3001 Leuven
Belgium
T: +32(0)475 2 8 7052
W: www.esat.kuleuven.be
E: bart.demoor@esat.kuleuven.be

A tribute

Venus in toys







Content

From least over total to structured total least squares

Linear matrix structure + rank deficiency = dynamical systems

The Riemannian singular value decomposition

Convexification of sets of multivariate polynomials

Examples

Least squares

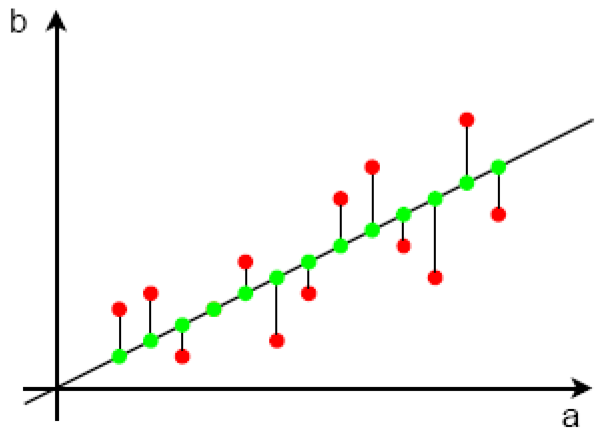
- Measurement:

$$Av \approx b$$

- Correction:

$$Av = b + \Delta b$$

- C.F. Gauss (± 1794): Predict future location of asteroid Ceres



$$\begin{array}{ll} \min & \|\Delta b\|_2^2\,, \\ \text{s.t.} & Av = b + \Delta b \end{array}$$

When Δb ~ N (0, σ I) → maximum likelihood

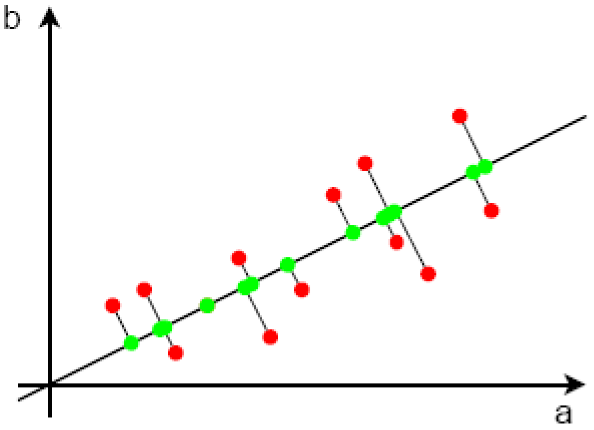
Total least squares

- Measurement:

$$Av \approx b$$

- Correction:

$$(A + \Delta A)v = (b + \Delta b)$$



$$\begin{array}{ll} \min & \left|\left|[\Delta A \quad \Delta b]\right|\right|_F^2 , \\ \text{s.t.} & (A + \Delta A)v = b + \Delta b, \\ & v^T v = 1 \end{array}$$

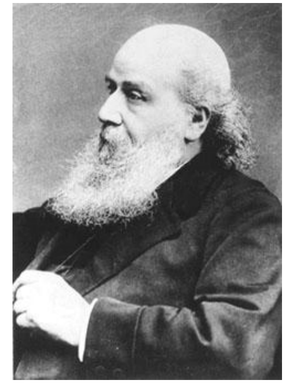
Solution = SVD of [A b]

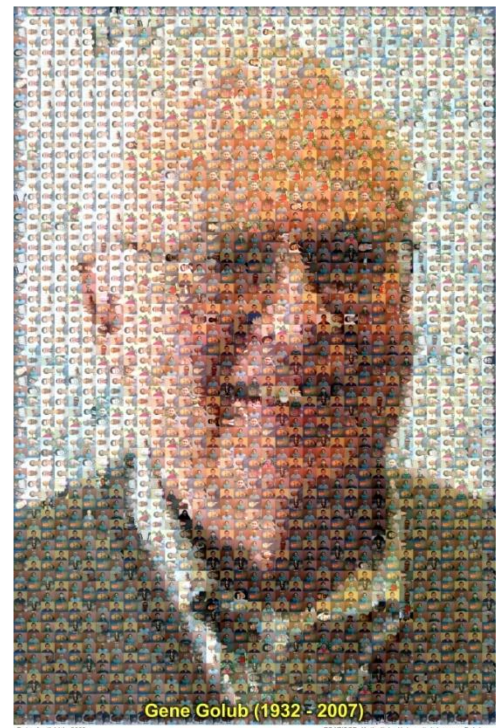
Long history:

Sylvester, Jordan, Schmidt, Eckart, Young, Mirsky, Golub

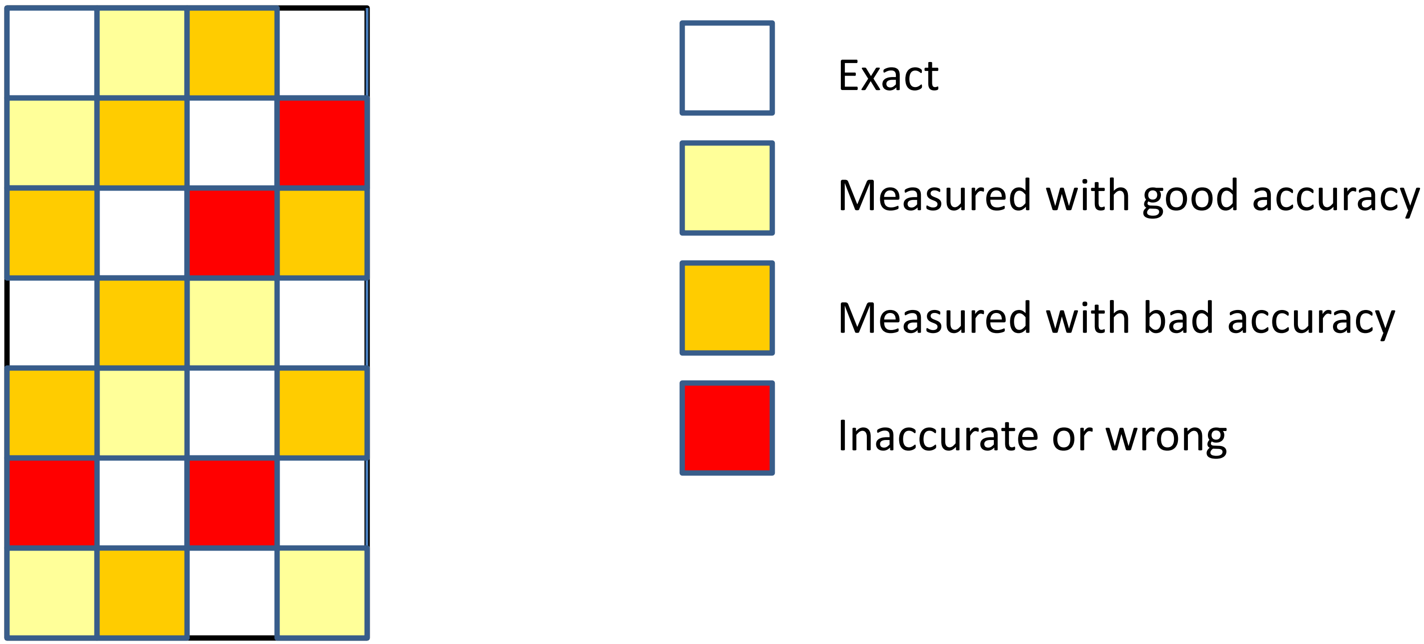
When noise Gaussian, maximum likelihood

When data Gaussian, relation to PCA





Weighted Total Least Squares



Geology: linear relations between mineral concentrarions
samples differ in size

Latent Semantic Analysis – rank reduction with sparse matrices

Structured Total Least Squares

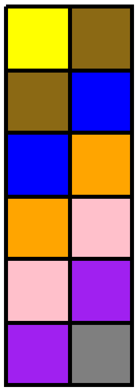
- Measurement:

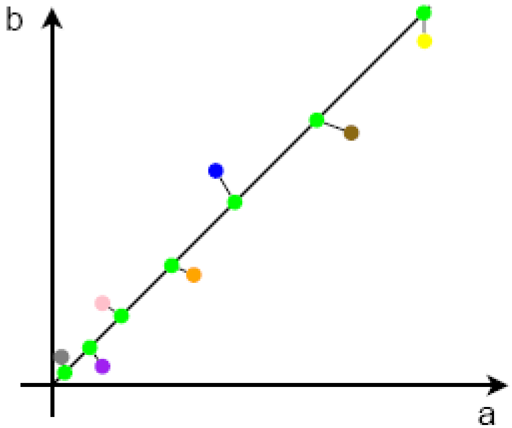
$$Av \approx b$$

- Correction:

$$(A + \Delta A)v = (b + \Delta b)$$

$$\begin{bmatrix}\Delta A & \Delta b\end{bmatrix} \text{ structured}$$





$$\begin{array}{ll} \min & \left\|\begin{bmatrix}\Delta A & \Delta b\end{bmatrix}\right\|_F^2, \\ \text{s.t.} & (A + \Delta A)v = b + \Delta b, \\ & v^T v = 1 \\ & \begin{bmatrix}\Delta A & \Delta b\end{bmatrix} \text{ structured} \end{array}$$

Linear matrix structure

+

Rank deficiency

=

Dynamical systems !

Realization theory

- Impulse response experiment: measure output data $h(k)$
- Construct Hankel matrix from data:

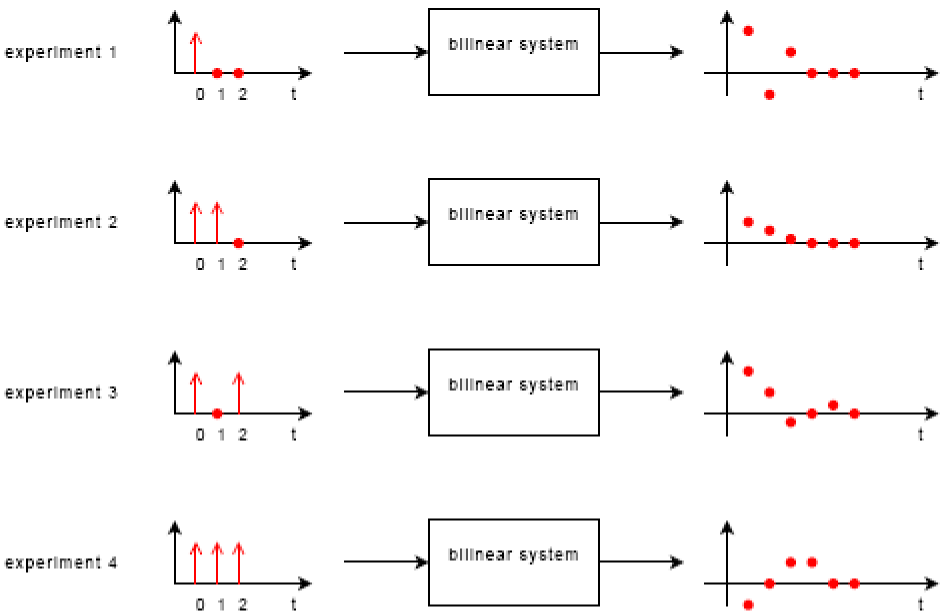
$$H = \begin{bmatrix} h(1) & h(2) & h(3) & h(4) & \dots \\ h(2) & h(3) & h(4) & & \\ h(3) & h(4) & & & \\ h(4) & & & & \\ \vdots & & & & \end{bmatrix}$$

- $\text{rank}(H) = \text{system order}$

Problem: SVD of Hankel ≠ Hankel !

Bilinear realization theory

$$
\begin{aligned}
x_{k+1} &= Ax_k + Bu_k + Nx_k \otimes u_k \\
y_k &= Cx_k + Du_k
\end{aligned}
$$



- **experiment 1:**
  $D,\ CB,\ CAB,\ CA^2B, \ldots$
- **experiment 2:**
  $D, CB + D,\ CAB + CNB + CB,$
  $CA^2B + CANB + CAB, \ldots$
- **experiment 3:**
  $D,\ CB,\ CAB,\ CA^2B,$
  $CA^3B + CNA^2B + CB, \ldots$
- **experiment 4:**
  $D,\ CB + D,$
  $CAB + CNB + CB + D,$
  $CA^2B + CANB + CAB +$
  $CNAB + CN^2B + CNB + CB, \ldots$

$$
\left[\begin{array}{c|cc|cccc}
CB & CAB & CNB & CA^2B & CANB & CNAB & CN^2B \\
\hline
CAB & CA^2B & CANB & CA^3B & CA^2NB & CANAB & CAN^2B \\
CNB & CNAB & CN^2B & CNA^2B & CNANB & CN^2AB & CN^3B \\
\hline
CA^2B & CA^3B & CA^2NB & CA^4B & CA^3NB & CA^2NAB & CA^2N^2B \\
CNAB & CNA^2B & CNANB & CNA^3B & CNA^2NB & CNANAB & CNAN^2B \\
CANB & CANAB & CAN^2B & CANA^2B & CANANB & CAN^2AB & CAN^3B \\
CN^2B & CN^2AB & CN^3B & CN^2A^2B & CN^2ANB & CN^3AB & CN^4B
\end{array}\right]
$$

$$
= \begin{bmatrix} C \\ CA \\ CN \\ CA^2 \\ CNA \\ CAN \\ CN^2 \\ \vdots \end{bmatrix} \begin{bmatrix} B & AB & NB & A^2B & ANB & NAB & N^2B & \dots \end{bmatrix}
$$

Linear matrix structure

+

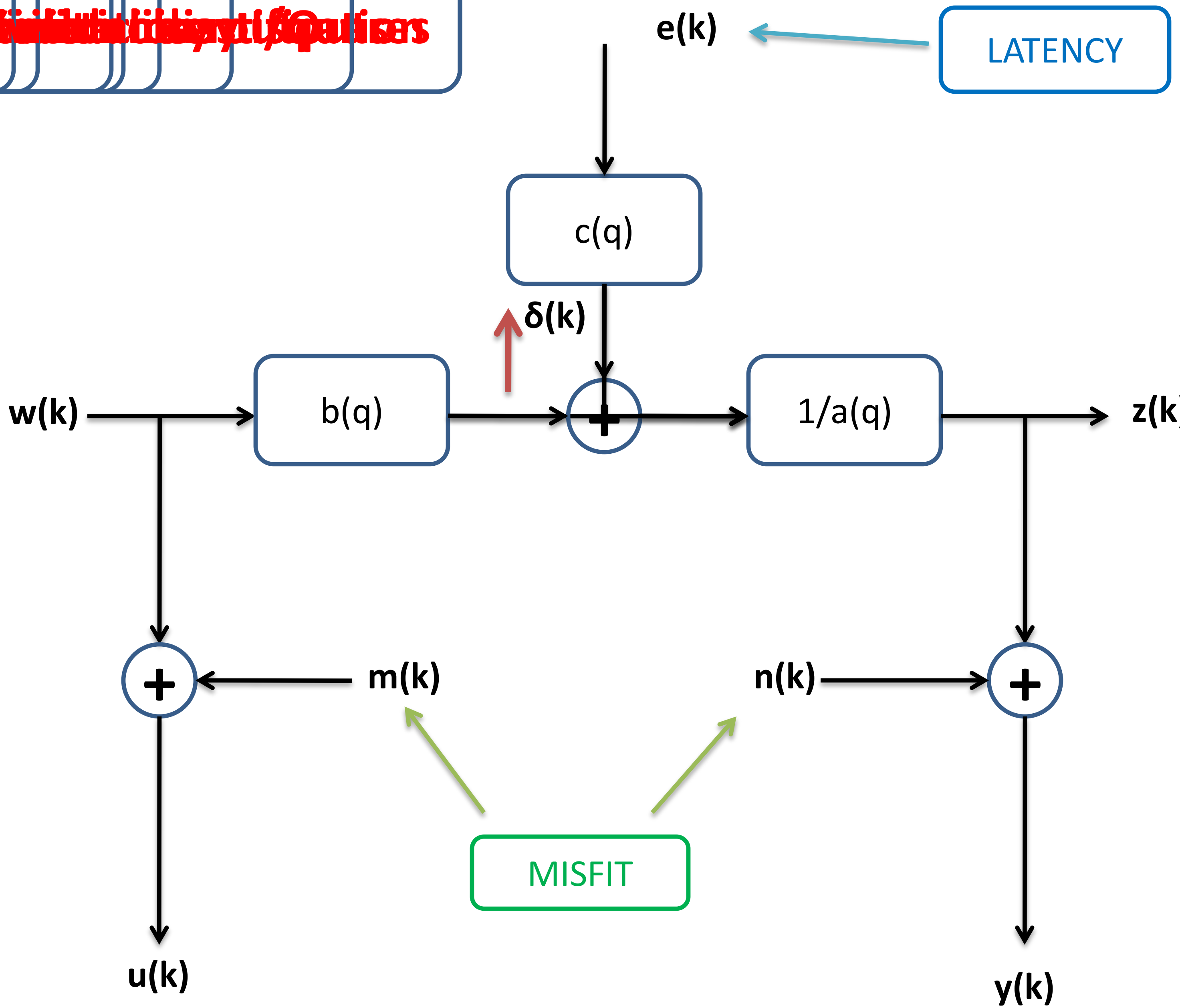
Rank deficiency

+

Noise on data

=

Dynamical system identification


e(k)
LATENCY
c(q)
δ(k)
w(k)
b(q)
1/a(q)
z(k)
m(k)
n(k)
MISFIT
u(k)
y(k)

(SISO) system identification problems

Minimize the sum of squares in the ‘misfit’ and the ‘latency’ subject to the fact that the ‘cleaned-up’ data should be related to each other by a linear time invariant dynamical system (= linearly structured matrix + rank deficiency)

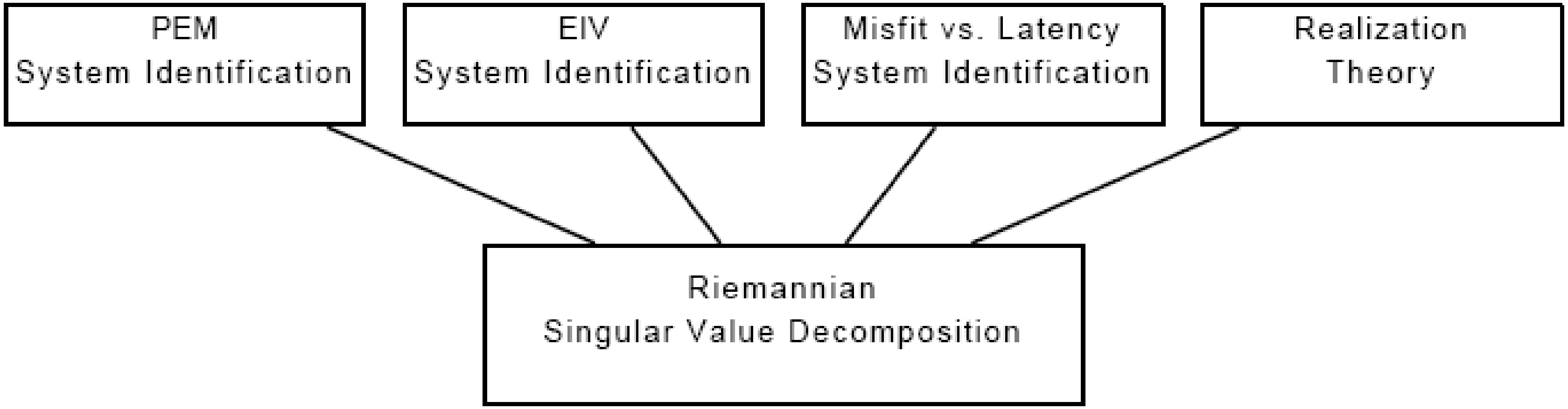
$J_{\text{output}} = \sum(\tilde{y}(t))^2$ $J_{\text{input}} = \sum(\tilde{u}(t))^2$ $J_{\text{latency}} = \sum(e(t))^2$

Given weights

$$\begin{array}{ll} \min & \alpha J_{\text{output}} + \beta J_{\text{input}} + \gamma J_{\text{latency}} \\ \text{s.t.} & Za - Wb - Ec = 0 \end{array}$$

Normalization constraints on a, b, c give rank deficiency

Hankel matrices (number of columns = degree of polynomials)



Find triplet (u, τ, v) for mimimal τ in

$$
\begin{aligned}
A^T u &= D_u v \tau \\
A v &= D_v u \tau \\
v^T v &= 1 \\
u^T D_v u &= 1 \; (= v^T D_u v)
\end{aligned}
$$

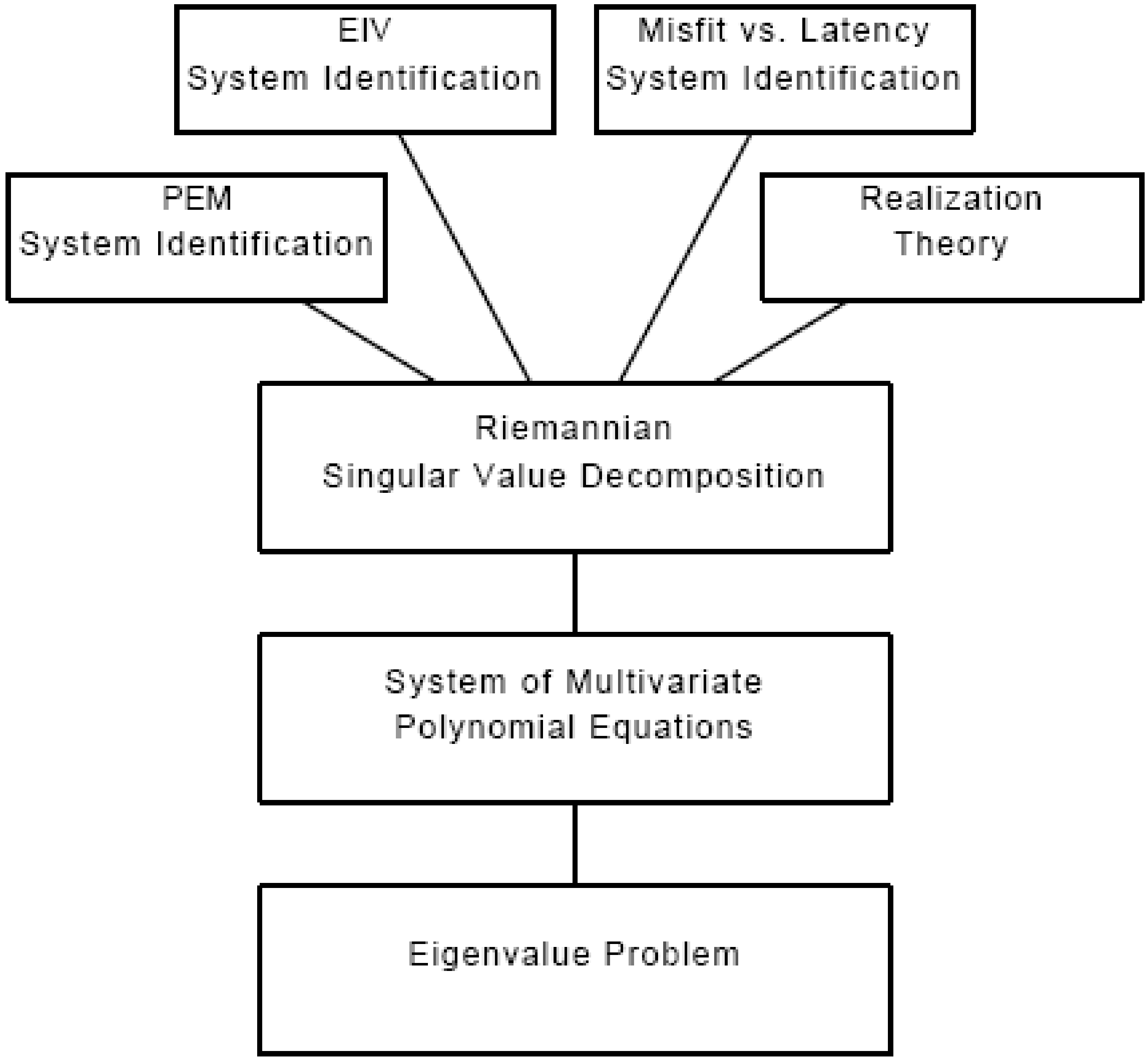
Structured data matrix

Positive definite quadratic matrix functions of u and v

Normalization constraints

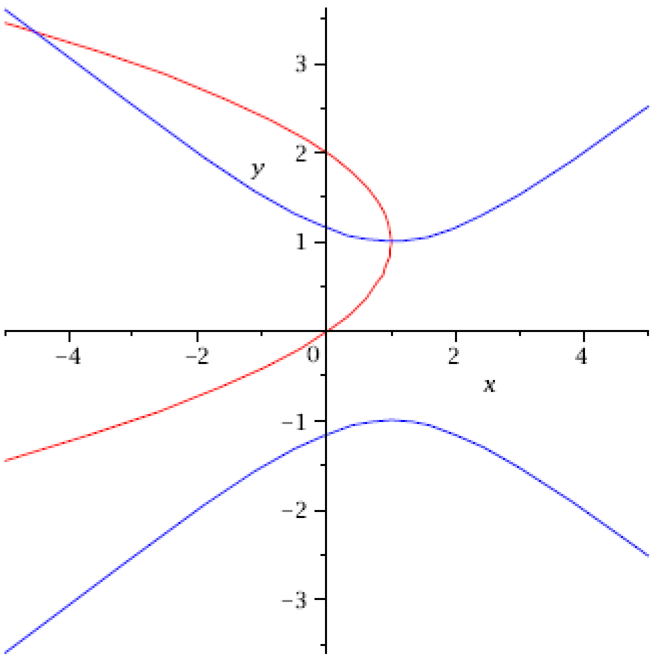
‘Nonlinear SVD’

= SET OF MULTIVARIATE POLYNOMIALS !!!

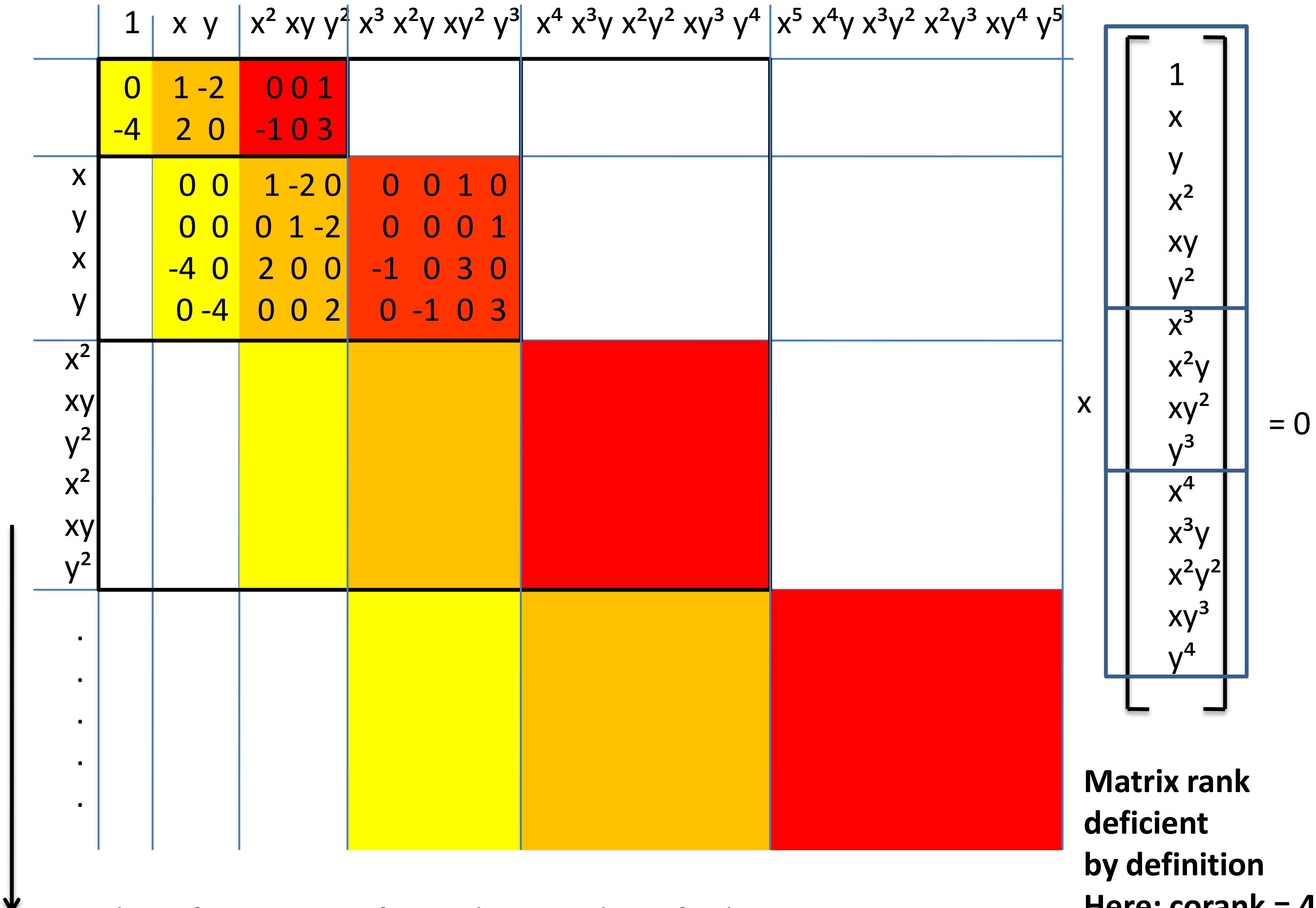

EIV
System Identification
Misfit vs. Latency
System Identification
PEM
System Identification
Realization
Theory
Riemannian
Singular Value Decomposition
System of Multivariate
Polynomial Equations
Eigenvalue Problem

Set of multivariate polynomials = eigenvalue problem(s)

$$y^2 - 2y + x = 0 \quad \text{(red)}$$
$$3y^2 - x^2 + 2x - 4 \quad \text{(blue)}$$



$y^2 - 2y + x = 0$
$3y^2 - x^2 + 2x - 4$



Number of rows grows faster than number of columns

Matrix rank deficient by definition
Here: corank = 4

Set of multivariate polynomials = eigenvalue problem(s)

$$y^2 - 2y + x = 0 \quad \text{(red)}$$
$$3y^2 - x^2 + 2x - 4 \quad \text{(blue)}$$

$$b = \begin{bmatrix} 1 & x & y & xy \end{bmatrix}^T$$

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ -4 & -1 & 6 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -6 & 8 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ xy \end{bmatrix} = x \begin{bmatrix} 1 \\ x \\ y \\ xy \end{bmatrix}$$

$A_1 \,.\, b = b \,.\, x$

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -1 & 2 & 0 \\ 4 & 1 & -6 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ xy \end{bmatrix} = y \begin{bmatrix} 1 \\ x \\ y \\ xy \end{bmatrix}$$

$A_2 \,.\, b = b \,.\, y$

Common eigenvectors → commuting matrices

But also.....

$$(A_1^2 + A_2^2)\ .\ b = b\ .\ (x^2+y^2)$$

↑

Can choose any
polynomial objective function
as an eigenvalue

Optimization problems with

polynomial objective functions

and polynomial constraints

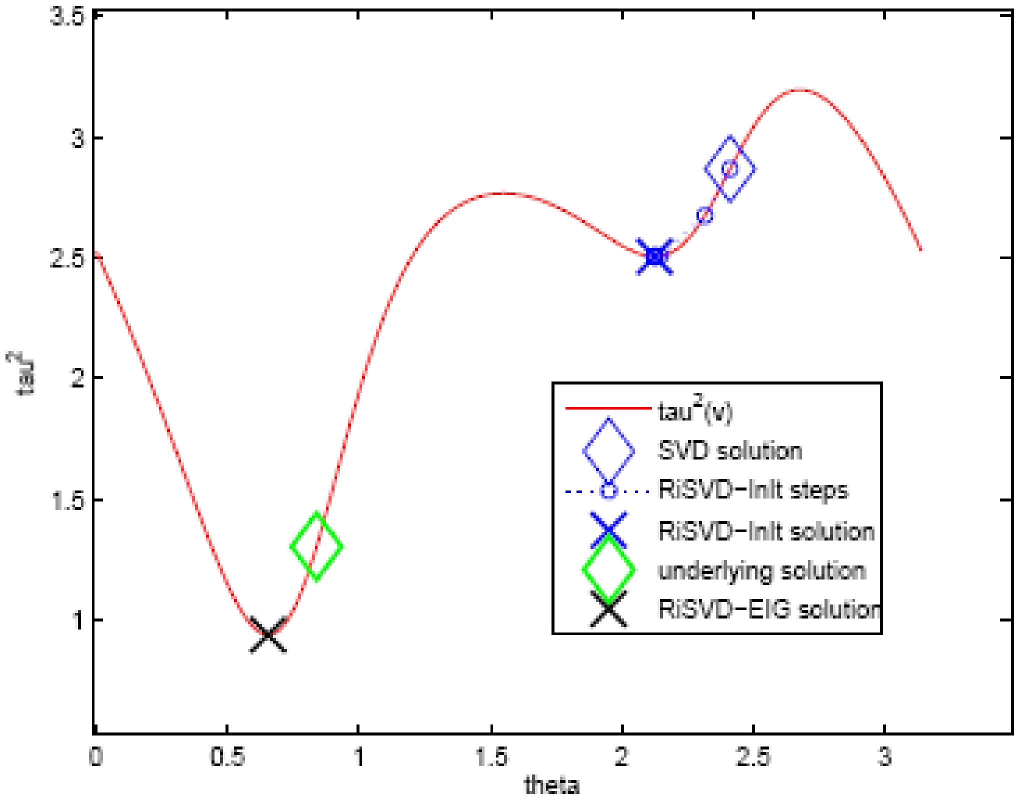
can always be written as

eigenvalue problems in which we look for the minimal eigenvalue

Convexification of polynomial optimization problems !!

3 x 2 Hankel

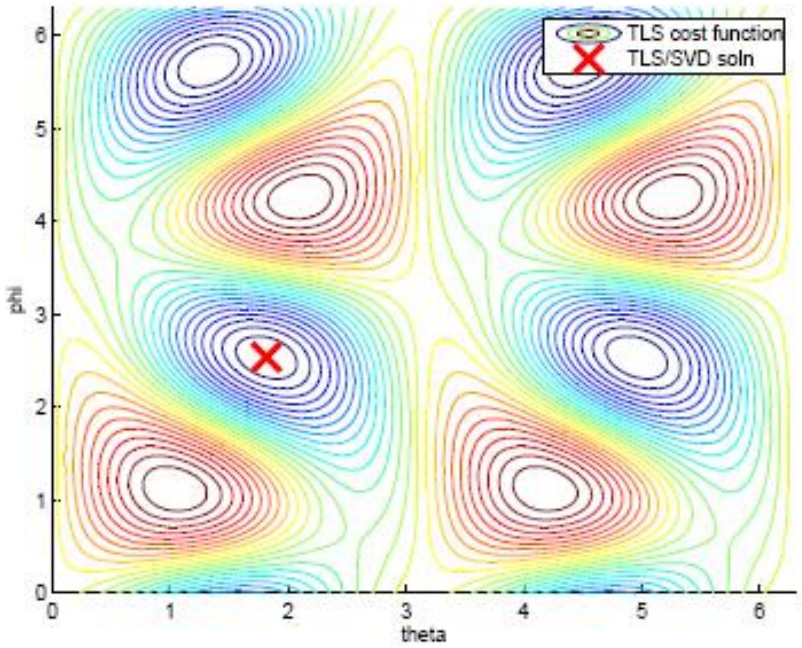
Least squares fit 5 data points
by impulse response
of first order system

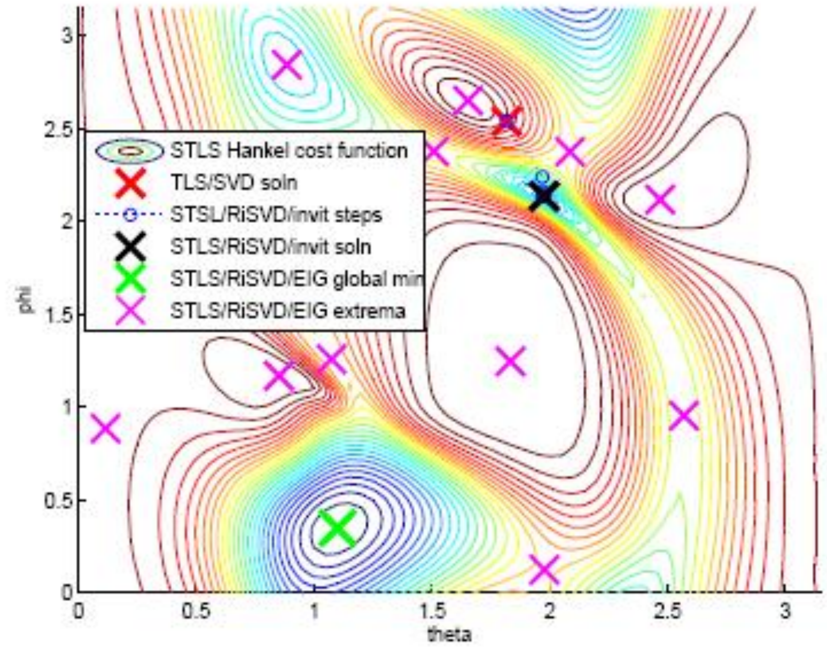


Eigenvalue decomposition on 20×20 matrix

6 x 3 Hankel

Least squares fit 8 data points
by impulse response
of second order system





Eigenvalue decomposition on 437×437 matrix

Conclusions

Riemannian SVD unifies

structured/weighted rank deficient least squares matrix approximations
PEM (prediction errors methods)
Many other approximation problems

Riemannian SVD is **finding 'minimizing' zero of set of multivariate polynomials**

Finding minimizing zero of set of multivariate polynomials is **extremal eigenvalue problem**

These relations in principle **'convexify' many identification problems** for LTI systems

Convexification occurs by **projecting up to higher dimensional vector space**
(difficult in low # dimensions; 'easy' in high # dimensions: an eigenvalue problem)

Can provide insight in other algorithms, yet not fully understood, such as subspace identification

Many challenges remain

- Effetient construction of the convexifying eigenvalue problem
- Algorithms to find the minimizing solution (inspired by inverse power method)
- Dealing with the curse of dimensionality
- How to cope with MIMO systems
- How to reduce the rank by more than 1
- etc….

At the end of day,
All we really understand,
Is linear algebra.

Adhemar,

Ad multos annos !!!