Moment Problems and Optimization: A Global-Analysis Approach to Problems in Systems and Control

Anders Lindquist
Shanghai Jiao Tong University, China, and the
Royal Institute of Technology, Stockholm, Sweden

International conference Optimization and Applications in Control and Data Science on the occasion of Boris Polyak's 80th birthday, May 13-15, 2015, Moscow, Russia

Dedicated to Boris Polyak at the occasion of his 80th birthday

The youngest man ever to turn 80

What is the talk about

- A classical problem – the moment problem – with a decidedly non-classical twist motivated by applications to systems and control.
- What is new are certain rationality constraints imposed by applications that alter the mathematical problem and make it nonlinear.
- A global-analysis approach that studies the class of solutions as a whole.
- A powerful paradigm for smoothly parametrizing, comparing, and shaping solutions to specifications.

Acknowledgements

Chris Byrnes

Tryphon Georgiou

Sergei Gusev

plus many others, among them Per Enqvist, Johan Karlsson, Ryozo Nagamune, Enrico Avventi, Axel Ringh, Giorgio Picci

The moment problem

Given $c_0, c_1, \ldots, c_n$, find $d\mu$ such that

$$\boxed{\int_a^b \alpha_k(t) d\mu = c_k, \quad k = 0, 1, \ldots, n}$$

$d\mu \in \mathcal{M}_+$ space of positive measures

- Power moment problem: $\alpha_k(t) = t^k$
- Trigonometric moment problem: $\alpha_k(t) = e^{ikt}, \quad [a, b] = [-\pi, \pi]$
- Nevanlinna-Pick interpolation: $\alpha_k(t) = \frac{e^{it}+z_k}{e^{it}-z_k}, \quad [a, b] = [-\pi, \pi]$

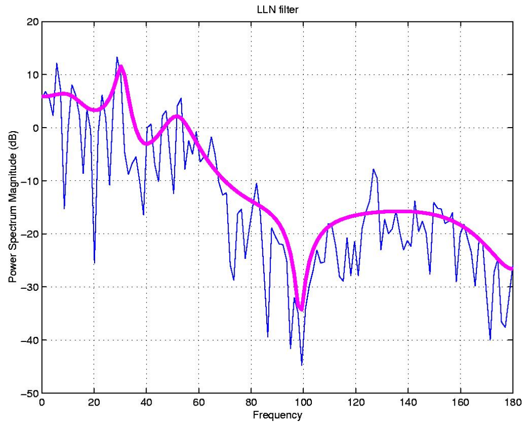
Chebyshev

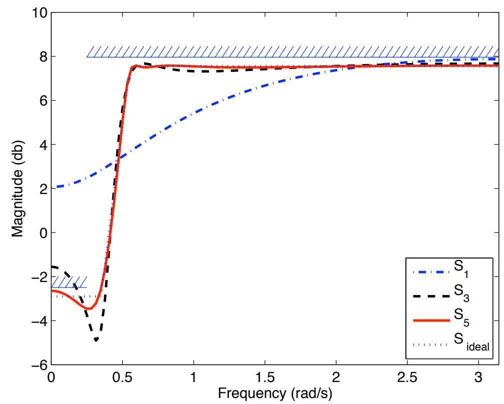
Markov

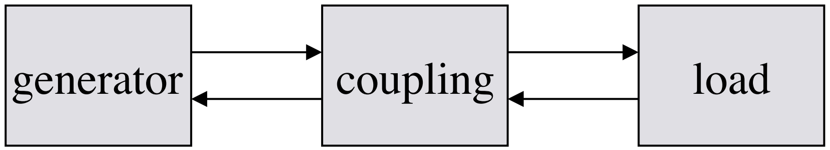
Lyapunov

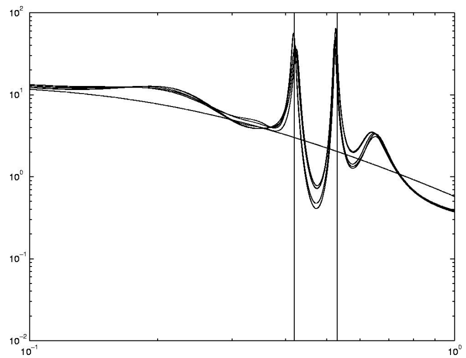
Where do we find moment problems in systems and control?

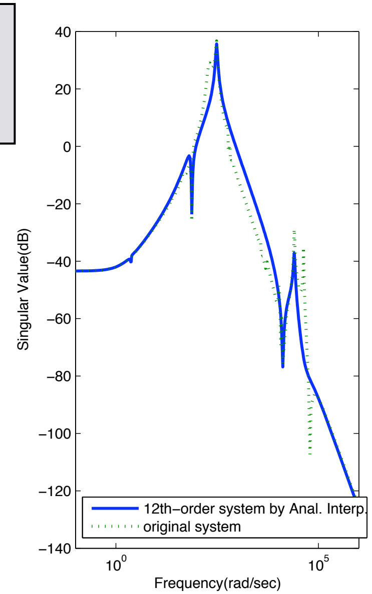
- spectral estimation
- speech synthesis
- system identification
- image processing
- optimal control
- robust control
- model reduction
- model matching problems
- simultaneous stabilization
- optimal power transfer











Multidimensional moment problems

$$\int_K \alpha_k d\mu = c_k, \quad k = 1, 2, \ldots, n$$

- $d\mu$ nonnegative measure on a compact subset K of $\mathbb{R}^d$
- $\alpha_1, \alpha_2, \ldots \alpha_n$ linearly independent basis functions defined on K

Image compression

Ringh, Karlsson and L.

… and why do these problems require a nonclassical approach?

- Solution must be of bounded complexity (such as, for example, rational of a bounded degree) reflecting the physical realizability of a finite-dimensional device
- Classical theory does not provide natural parameterizations of rational solutions of bounded degree

Moment problems come in many different forms

Let us look at a few and return to them throughout the lecture

Ex. 1: Covariance extension

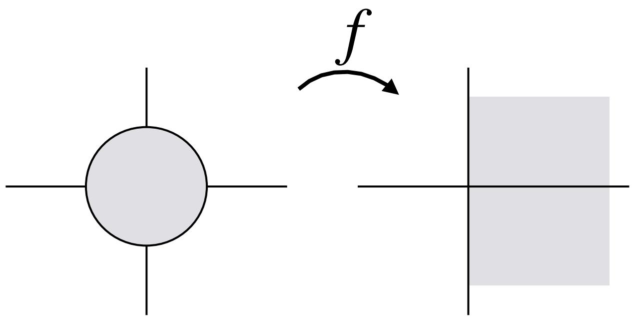
- $c_k = E\{y(t+k)y(t)\}$, where y stationary stochastic process

Given $c_0, c_1, \dots, c_n$, find an infinite extension $c_{n+1}, c_{n+2}, \dots$ such that

$$f(z) = \frac{1}{2}c_0 + c_1 z + \cdots + c_n z^n + c_{n+1} z^{n+1} + \dots$$

(i) is positive real (Cathéodory)

(ii) is rational of degree at most n



spectral estimation, speech synthesis, system identification

Covariance extension (alternative formulation)

PROBLEM: Given $c_0, c_1, \ldots, c_n$ such that

$$T_n = \begin{bmatrix} c_0 & c_1 & \cdots & c_n \\ \bar{c}_1 & c_0 & \cdots & c_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{c}_n & \bar{c}_{n-1} & \cdots & c_0 \end{bmatrix} > 0$$

Find infinite extension $c_{n+1}, c_{n+2}, \ldots$ such that

$$T_\infty = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots \\ \bar{c}_1 & c_0 & c_1 & \cdots \\ \bar{c}_2 & \bar{c}_1 & c_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} > 0$$

(ii) is rational of degree at most n

Ex 2: Circulant covariance extension

- $c_k = E\{y(t+k)y(t)\}$, where y periodic stochastic process

$$\boxed{c_{2N-k} = c_k, \quad k = 0, 1, \ldots, N}$$

reciprocal process
(Jamison, Krener, Levy, Frezza, Ferrante, Picci, Pavon, Carli)

PROBLEM: Given $n < N$ and $c_0, c_1, \ldots, c_n$ such that $T_n > 0$, find extension $c_{n+1}, c_{n+2}, \ldots, c_N$ such that

$$T_{2N-1} = \begin{bmatrix} c_0 & c_1 & \cdots & c_1 \\ \bar{c}_1 & c_0 & \cdots & c_2 \\ \vdots & \vdots & \ddots & \vdots \\ \bar{c}_1 & \bar{c}_2 & \cdots & c_0 \end{bmatrix} > 0$$

circulant matix

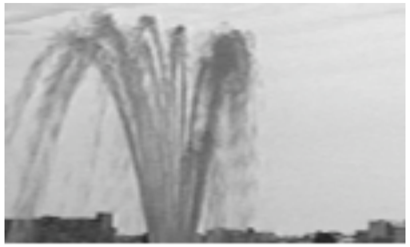


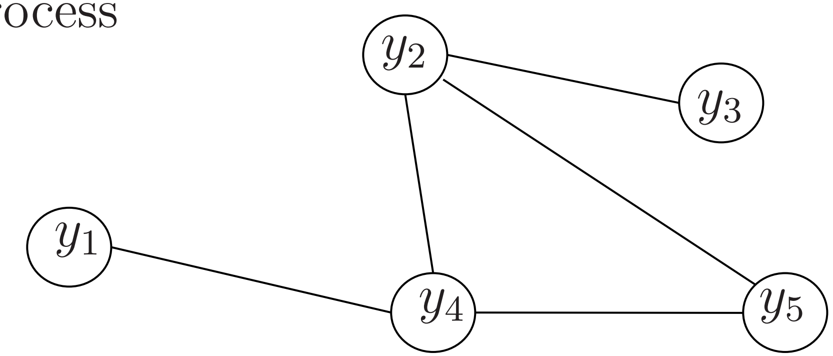
Image processing (in vector case)

(Carli, Ferrante, Pavon, Picci)

Ex 3: Identification of graphical models

$\{y(t)\}_{t\in\mathbb{Z}}$ m-dimensional stationary, Gaussian stochastic process

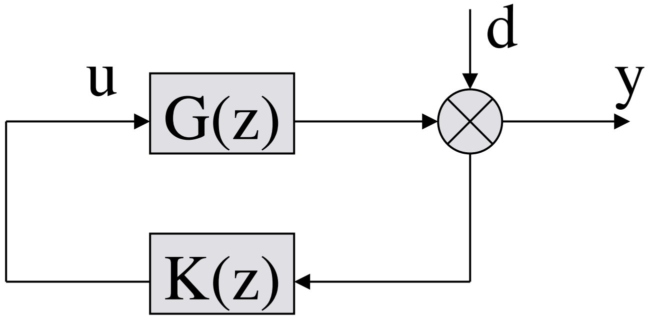
The components of y are represented as nodes in a graph, where the lack of edge between y_k and y_ℓ implies that y_k and y_ℓ are conditionally independent given the rest of the components.



PROBLEM 1. Given $m \times m$ covariance lags $c_0, c_1, \ldots, c_n$, find an ARMA model for y respecting the conditional independence.

PROBLEM 2. Given observed sample data from y, determine the topology of the graph.

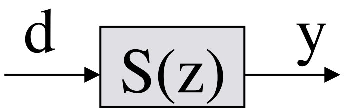
Ex. 4: Robust control



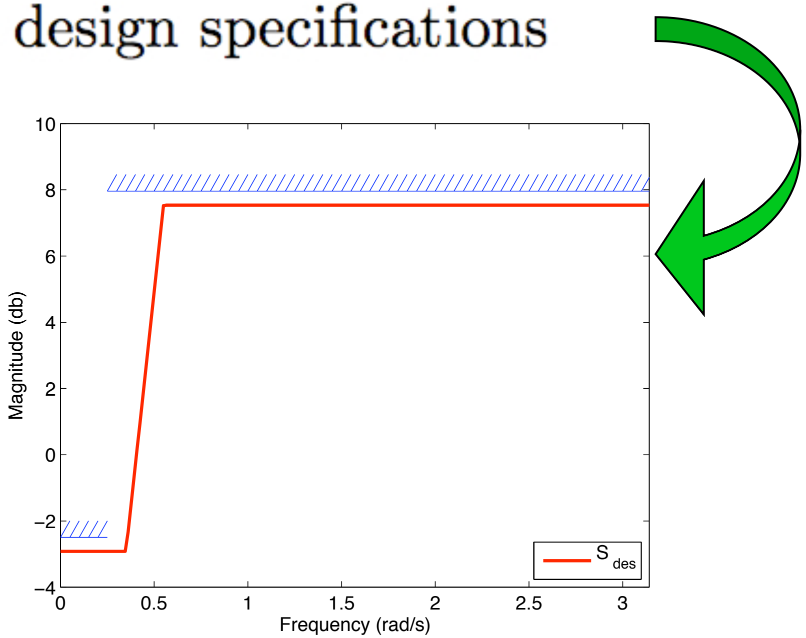
PROBLEM: Find an internally stabilizing controller K such that

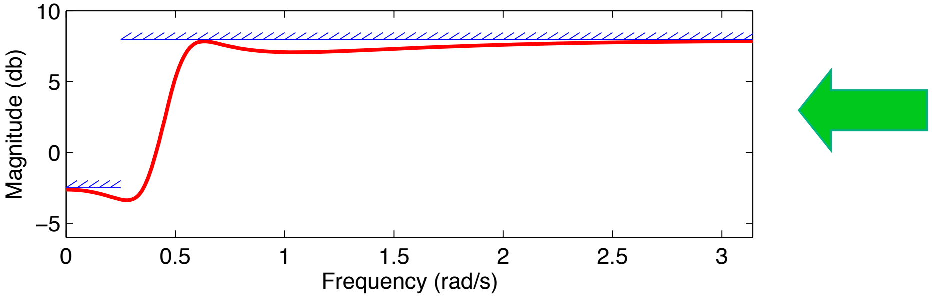
$$S = (I - GK)^{-1}$$

has low degree and satisfies the design specifications



sensitivity function





What do we want to achieve?

- A basic paradigm for smooth parameterization of the whole class of solutions in a systems-theoretical language
- Methods for determining solutions by convex optimization

Moment problem in the style of Krein

$\mathfrak{P}$ finite-dimensional subspace of $C[a,b]$

$(\alpha_0, \alpha_1, \ldots, \alpha_n)$ basis in $\mathfrak{P}$

Given $c := (c_0, c_1, \ldots, c_n) \in \mathbb{C}^{n+1}$,

find positive measure $d\mu$ such that

$$\int_a^b \alpha_k(t) d\mu = c_k, \quad k = 0, 1, \ldots, n$$

$$\mathfrak{P}_+ := \{p \in \mathfrak{P} \mid P(t) := \mathrm{Re}\{p(t)\} \geq 0 \quad \forall t \in [a,b]\}$$

positive cone

closed convex

Suppose $\mathfrak{P}_+$ has a nonempty interior $\mathring{\mathfrak{P}}_+$

Dual cones

$$\mathfrak{P}_+ := \left\{ p = \sum_{k=0}^{n} p_k \alpha_k \mid P(t) := \mathrm{Re}\{p(t)\} \geq 0 \quad \forall t \in [a,b] \right\}$$

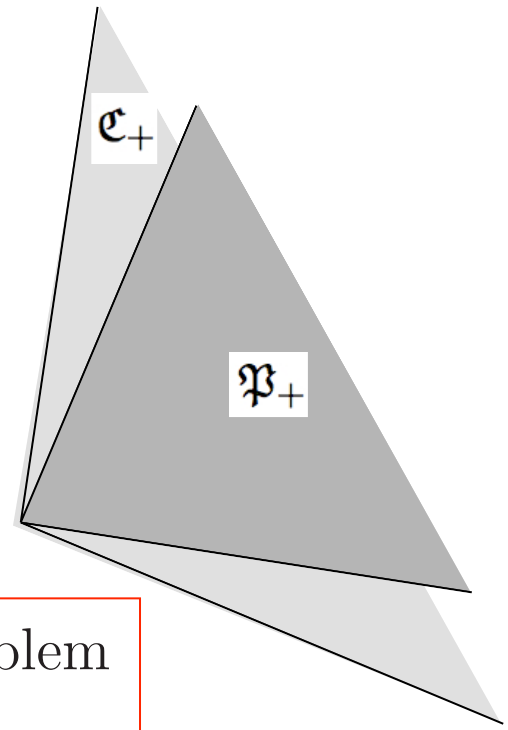
closed convex cone

For any $c = (c_0, c_1, \ldots, c_n) \in \mathbb{C}^{n+1}$, form

$$\langle c, p \rangle := \mathrm{Re} \left\{ \sum_{k=0}^{n} c_k p_k \right\}$$

$$\mathfrak{C}_+ := \{ c \in \mathbb{C}^{n+1} \mid \langle c, p \rangle \geq 0 \quad \forall p \in \mathfrak{P}_+ \}$$

dual cone, closed convex



THEOREM(Krein-Nudelman) The moment problem

$$\int_a^b \alpha_k(t) d\mu(t) = c_k, \quad k = 0, 1, \ldots, n$$

is solvable if and only if $c \in \mathfrak{C}_+$.

$$\boxed{\text{THEOREM (Krein-Nudelman) The moment problem } \int_a^b \alpha_k(t)d\mu(t) = c_k, \quad k = 0, 1, \dots, n \text{ is solvable if and only if } c \in \mathfrak{C}_+.}$$

THEOREM(Krein-Nudelman) The moment problem

$$\int_a^b \alpha_k(t)d\mu(t) = c_k, \quad k = 0, 1, \dots, n$$

is solvable if and only if $c \in \mathfrak{C}_+$.

Trigonometric moment problem = covariance extension problem

$$\boxed{\alpha_k(t) = e^{ikt}} \qquad c \in \mathfrak{C}_+ \iff T_n = [c_{k-\ell}]_{k,\ell=0}^n \geq 0$$

Toeplitz matrix

Nevanlinna-Pick interpolation

$$\boxed{\alpha_k(t) = \frac{e^{it} + z_k}{e^{it} - z_k}} \qquad c \in \mathfrak{C}_+ \iff P_n = \left[\frac{c_j + \bar{c}_k}{1 - z_j \bar{z}_k}\right]_{j,k=0}^n \geq 0$$

Pick matrix

Introducing a finiteness condition

Consider the class $\mathcal{R}_+$ of rational positive measures

$$d\mu = \frac{P(t)}{Q(t)} dt$$

$$\boxed{\begin{array}{l} p(t) = \sum_{k=0}^{n} p_k \alpha_k(t) \\ \text{polynomial in } \mathfrak{P}_+ \end{array}}$$

where $P(t) = \mathrm{Re}\{p(t)\}, Q(t) = \mathrm{Re}\{q(t)\}$ and $p, q \in \mathring{\mathfrak{P}}_+$.

$\mathcal{R}_+ \subset \mathcal{M}_+$ the space of positive measures

$$\int_a^b \alpha_k d\mu = c_k, \quad k = 0, 1, \ldots, n \quad (\dagger)$$

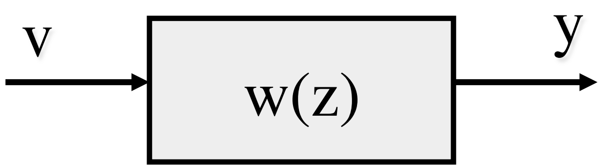
Find $d\mu \in \mathcal{M}_+$ satisfying $(\dagger)$ linear problem

Find $d\mu \in \mathcal{R}_+$ satisfying $(\dagger)$ nonlinear problem

Nevertheless, our approach is based entirely on convexity.

A motivating example

white noise



stationary process
with spectral density
$\Phi(e^{i\theta}) = |w(e^{i\theta})|^2$

$w(z)$ rational transfer function of degree n
with no poles and zeros on the unit circle

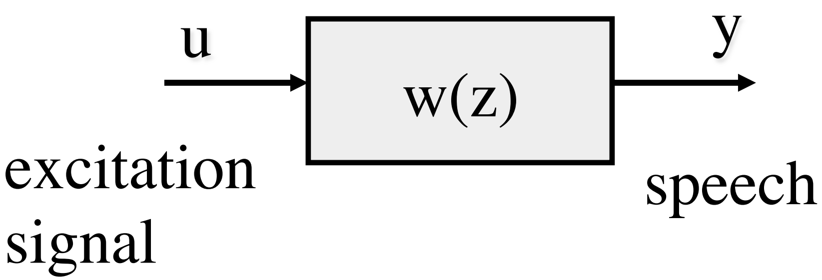
$$\int_{-\pi}^{\pi} e^{ik\theta}\Phi(e^{i\theta})\frac{d\theta}{2\pi} = c_k := E\{y(t+k)y(t)'\}$$

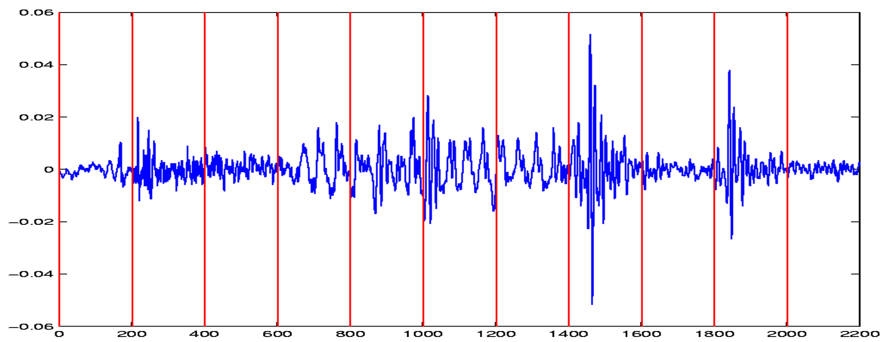
$$\boxed{\frac{1}{2\pi}\Phi(e^{i\theta}) = \frac{P(\theta)}{Q(\theta)}}$$

$\mathfrak{P}$ consists of trigonometric polynomials

$$P(\theta) = \mathrm{Re}\{p(\theta)\}, \quad Q(\theta) = \mathrm{Re}\{q(\theta)\}, \text{ where } p, q \in \mathring{\mathfrak{P}}_+$$

Ex: Modeling speech





on each (30 ms) subinterval
w(z) constant, y stationary

observation: $y_0, y_1, \ldots, y_N$

$N \approx 250$

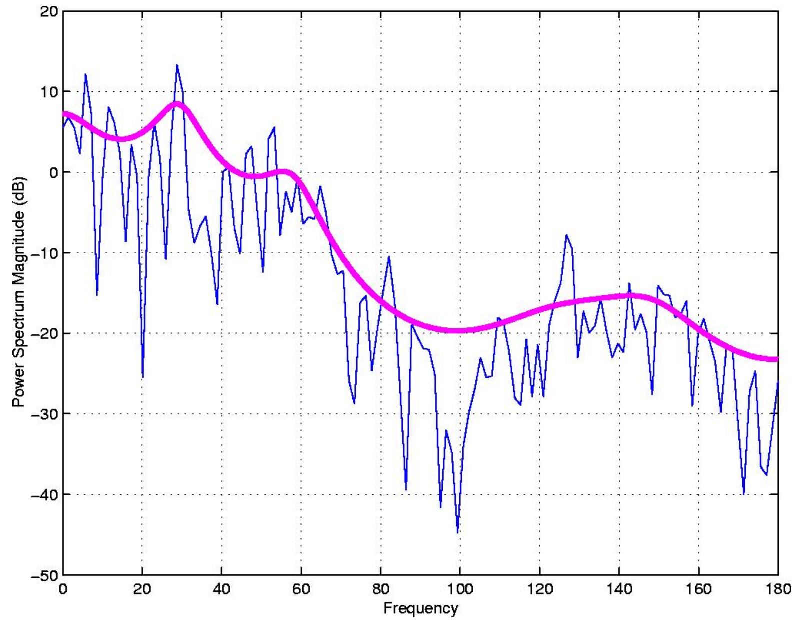
$$\int_{-\pi}^{\pi} e^{ikt} d\mu = c_k := \frac{1}{N+1} \sum_{t=0}^{N-k} y_{t+k} y_t, \quad k = 0, 1, \ldots, n \quad n = 10$$

$$\mathfrak{P} = \operatorname{span}\{1, e^{it}, e^{2it}, \ldots, e^{int}\}$$

$$\boxed{d\mu = \left|w(e^{it})\right|^2 dt \in \mathcal{R}_+}$$

Cellular telephone:

$$d\mu = \frac{\rho_n}{|\varphi_n(e^{it})|^2} dt$$



FFT in blue envelope in purple

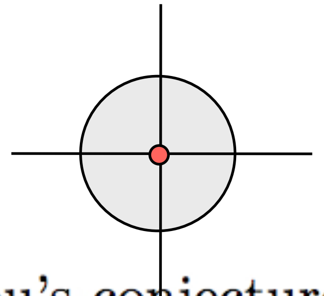
$\varphi_n(z)$ n:th Szegö polynomial orthogonal on the unit circle

$$P(t) = \rho_n \qquad Q(t) = \left|\varphi_n(e^{it})\right|^2$$

$\sigma(z)$ stable polynomial spectral factor: $\left|\sigma(e^{it})\right|^2 = P(t)$

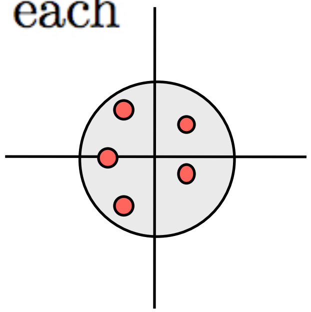
$$\sigma(z) = \sqrt{\rho_n} z^n$$

spectral zeros



Is there a solution $d\mu$ for each choice of spectral zeros?

YES (Georgiou 1983)

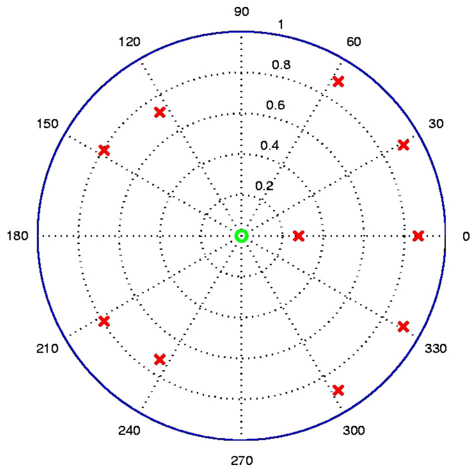


Unique? (Georgiou's conjecture)

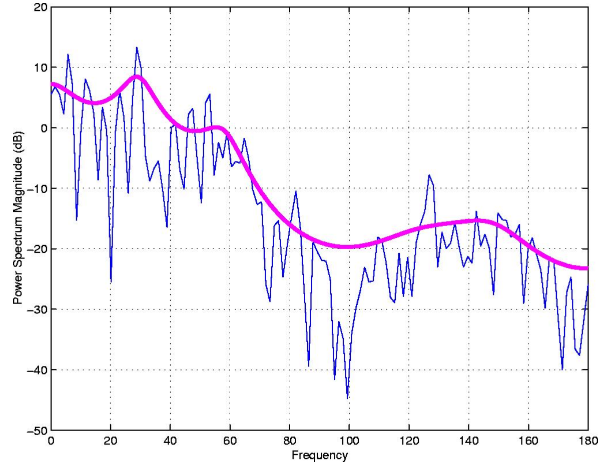
Well-posed?

YES (Byrnes, Lindquist Gusev, Matveev 1993)

$$w(z) = \frac{\sqrt{\rho_n} z^n}{\varphi_n(e^{it})}$$

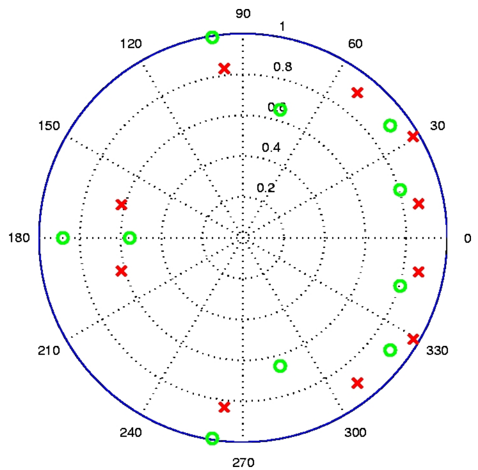


zeros/poles

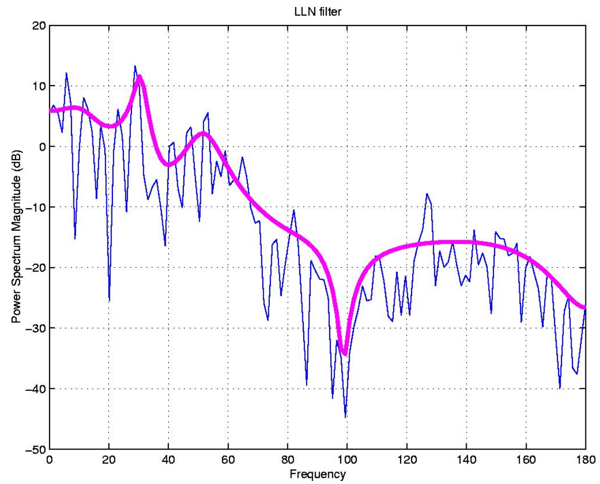


envelope

A $w(z)$ with other spectral zeros, but with the same degree



zeros/poles



envelope

THEOREM The moment problem for $\mathcal{R}_+$ measures

$$\int_a^b \alpha_k(t)\frac{P(t)}{Q(t)}dt = c_k, \quad k = 0, 1, \ldots, n$$

is solvable if and only if $c \in \mathring{\mathfrak{C}}_+$, the interior of $\mathfrak{C}_+$.

Compare with the classical problem for $\mathcal{M}_+$ measures:

THEOREM(Krein-Nudelman) The moment problem

$$\int_a^b \alpha_k(t)d\mu(t) = c_k, \quad k = 0, 1, \ldots, n$$

is solvable if and only if $c \in \mathfrak{C}_+$.

THEOREM The moment problem for $\mathcal{R}_+$ measures

$$\int_a^b \alpha_k(t) \frac{P(t)}{Q(t)} dt = c_k, \quad k = 0, 1, \ldots, n$$

is solvable if and only if $c \in \mathring{\mathfrak{C}}_+$, the interior of $\mathfrak{C}_+$.

In fact, if $c \in \mathring{\mathfrak{C}}_+$, the moment problem for $\mathcal{R}_+$ measures with $p \in \mathring{\mathfrak{P}}_+$ fixed is solvable.

THEOREM For any $(c, p) \in \mathring{\mathfrak{C}}_+ \times \mathring{\mathfrak{P}}_+$, there exists a $q \in \mathring{\mathfrak{P}}_+$ such that

$$\int_a^b \alpha_k(t) \frac{P(t)}{Q(t)} dt = c_k, \quad k = 0, 1, \ldots, n.$$

We want to show that there is a unique solution for each $p \in \mathring{\mathfrak{P}}_+$.

A Dirichlet principle

$$\boxed{\mathbb{J}_p(q) = \langle c, q \rangle - \int_a^b P \log Q \, dt}$$

strictly convex function

$\mathbb{J}_p : \mathring{\mathfrak{P}}_+ \to \mathbb{R}$

Moment equations:

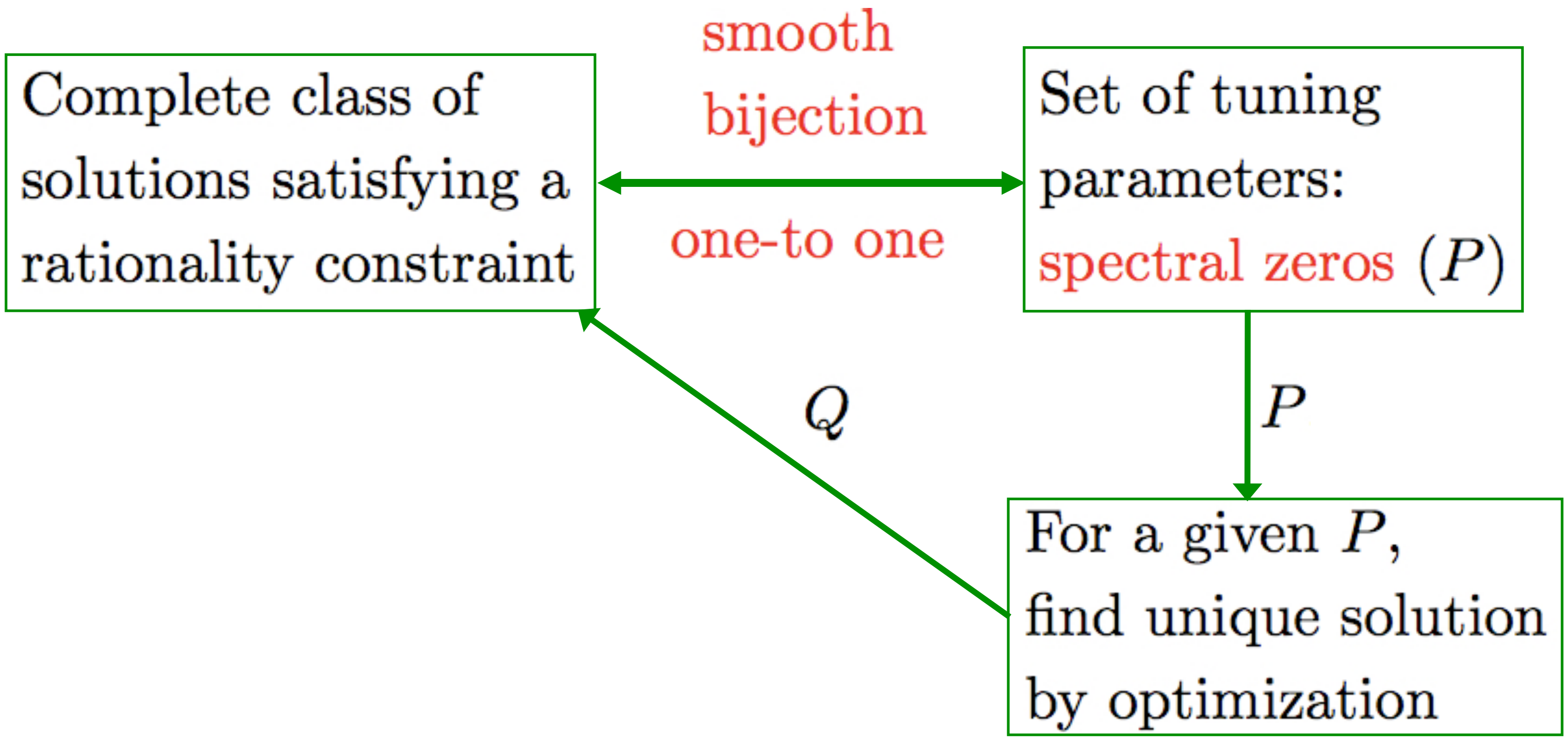
$$\frac{\partial \mathbb{J}_p}{\partial q_k} = c_k - \int_a^b \alpha_k \frac{P}{Q} dt = 0, \quad k = 0, 1, \ldots n \quad (\dagger)$$

We have already pointed out that the moment equations $(\dagger)$ have a solution $\hat{q} \in \mathring{\mathfrak{P}}_+$ for all $(c, p) \in \mathring{\mathfrak{C}}_+ \times \mathring{\mathfrak{P}}_+$. Since $\mathbb{J}_p$ is strictly convex, $\hat{q}$ is a unique minimum. Hence $(\dagger)$ has a unique solution.

THEOREM Let $(c, p) \in \mathring{\mathfrak{C}}_+ \times \mathring{\mathfrak{P}}_+$, and set $P := \mathrm{Re}\{p\}$. Then the functional $\mathbb{J}_p$ has a unique minimizer $\hat{q} \in \mathring{\mathfrak{P}}_+$, and $\hat{Q} := \mathrm{Re}\{\hat{q}\}$ is the unique solution of the moment equations $(\dagger)$.

A global analysis approach

Object: Finding a solution that best satisfies additional design specifications (without increasing the complexity)

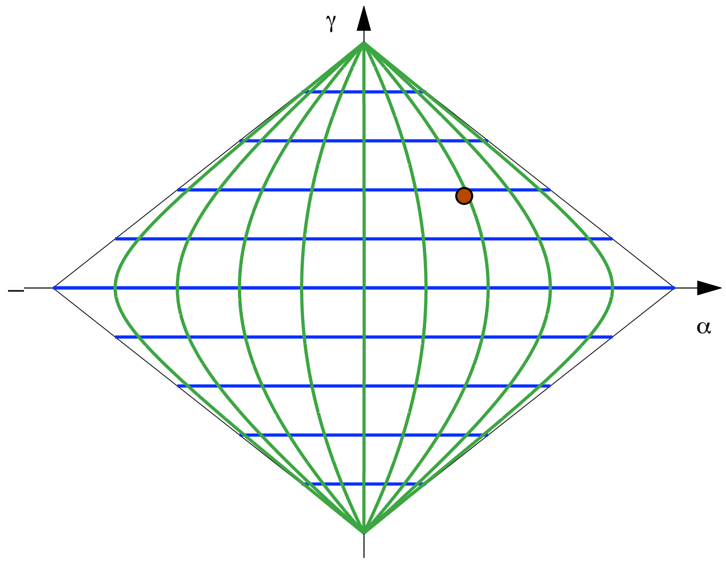


EXAMPLE. $\mathfrak{P} = \text{span}\{1, e^{it}, \dots, e^{int}\}$

The solutions $d\mu \in \mathcal{R}_+$ form a manifold of dimension $2n$.

A foliation with one leaf for each choice of $p \in \mathring{\mathfrak{P}}_+$ (Kalman filtering)

A foliation with one leaf for each choice of $c \in \mathring{\mathfrak{C}}_+$ (normalized)



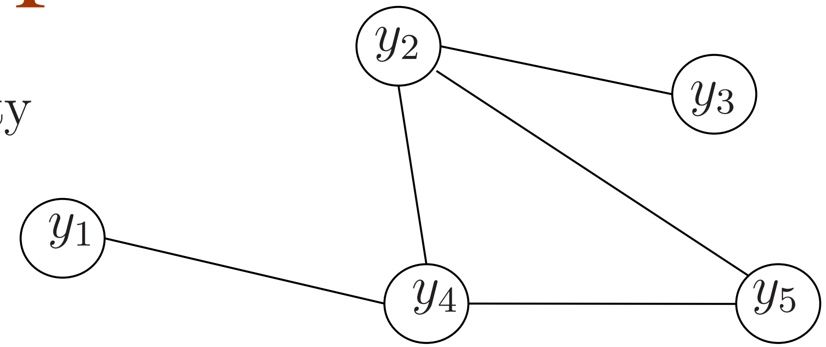
THEOREM. The two foliations intersect transversely so that each leaf in one meets each leaf in the other in exactly one point.

$$\min_{q \in \mathring{\mathfrak{P}}_+} \mathbb{J}(q)$$

unique solution $d\mu = \frac{P}{Q} dt$

Identification of graphical models

$\{y(t)\}_{t\in\mathbb{Z}}$ $\quad\Phi\;\; m\times m$ spectral density



$$(\mathbf{D})\quad \mathbb{J}_p(Q)=\langle c,q\rangle-\int_{-\pi}^{\pi}P\log\det Q dt\rightarrow\min$$

subject to $\quad Q_{k\ell}\equiv 0,\quad (y_k,y_\ell)\in\overline{\{\text{edges}\}}$ $\quad$ Avventi–L–Wahlberg

$$\Phi=\frac{P}{Q}\qquad \left[\Phi^{-1}\right]_{k\ell}\equiv 0$$

y_k and y_ℓ conditionally independent given the rest

$$(\mathbf{D}')\quad \mathbb{J}_p(Q)+\lambda\|Q\|_1\rightarrow\min$$

Cf. Songsiri –Vandenberghe for the special case $P\equiv 1$

Circulant covariance extension

$$c_{2N-k} = c_k, \quad k \leq N$$

$$\mathfrak{P}_+(N) = \{p \in \mathfrak{P} \mid P(e^{ik\pi/N}) > 0, \; k = 0, 1, \dots, 2N\} \supset \mathfrak{P}_+$$

$$\mathfrak{C}_+(N) = \mathfrak{P}_+(N)^{\mathsf{T}} \subset \mathfrak{C}_+$$

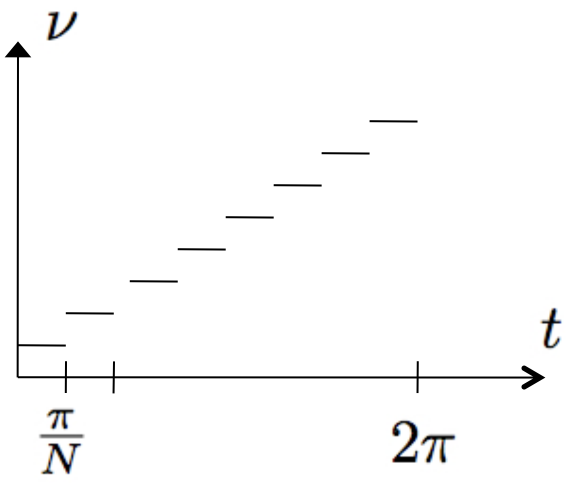
$$\mathfrak{C}_+(N) \to \mathfrak{C}_+ \text{ as } N \to \infty$$

$$\mathcal{R}(N) = \{d\mu \in \mathcal{M}_+ \mid d\mu = \frac{P}{Q} d\nu, \; p, q \in \mathfrak{P}_+(N)$$

sum of Dirac measures

PROBLEM. Given $c_0, c_1, \dots, c_n$ $(n < N)$ find $d\mu \in \mathcal{R}(N)$ such that

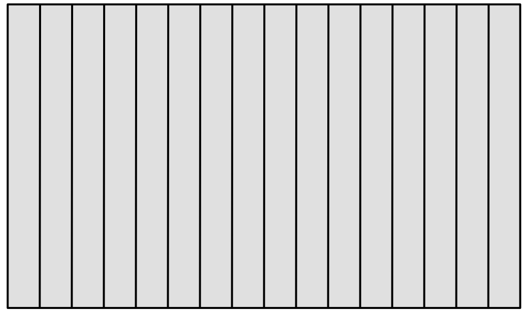
$$\int_{-\pi}^{\pi} e^{ikt} d\mu = c_k, \; k = 0, 1, \dots, n \quad (\dagger)$$



For each $(c, p) \in \mathfrak{C}_+(N) \times \mathfrak{P}_+(N)$, there is a unique $q \in \mathfrak{P}_+(N)$ such that $(\dagger)$ holds. It is the unique minimizer of

$$\mathbb{J}_p(q) = \langle c, q \rangle - \int_{-\pi}^{\pi} P \log Q \, d\nu,$$

Image processing



$y(t)$ reciprocal m-vector process

$$c_k = E\{y(t+k)y(t)^{\mathsf{T}}\} \quad m \times m$$

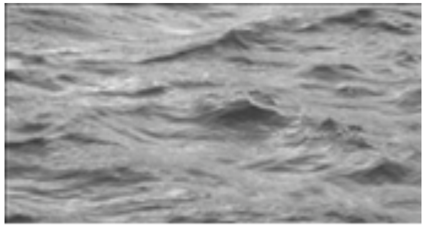
$$c_{2N-k} = {c_k}^{\mathsf{T}}, \quad k \leq N$$

For scalar P there is a matrix version of

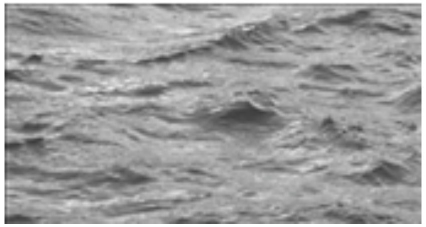
$$\boxed{\mathbb{J}_p(q) = \langle c, q \rangle - \int_{-\pi}^{\pi} P \log Q \, d\nu,}$$

original below

Francesca Carli, Augusto Ferrante, Michele Pavon, and Giorgio Picci



reconstructions with $P = 1$ (maximum entropy) and $n = 1$ $(m = 125,\ 2N - 1 = 175)$



Nevanlinna-Pick interpolation

$$\alpha_k(t) = \frac{e^{it} + z_k}{e^{it} - z_k}, \quad k = 0, 1, \ldots, n \qquad z_0, z_1, \ldots, z_n \in \mathbb{D}$$

(distinct)

$$c \in \mathfrak{C}_+ \quad \Longleftrightarrow \quad P_n = \left[\frac{c_j + \bar{c}_k}{1 - z_j \bar{z}_k} \right]_{j,k=0}^{n} \geq 0$$

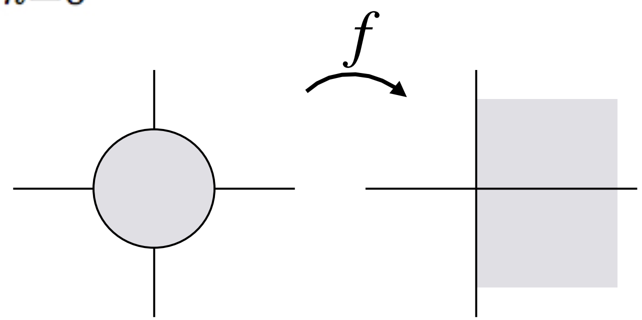
Pick matrix

Given $z_0, z_1, \ldots, z_n \in \mathbb{D}$ (distinct), find a Carathéodory function f such that

$$\boxed{f(z_k) = c_k, \quad k = 0, 1, \ldots, n}$$

$$\Big\Updownarrow$$

$$\int_{-\pi}^{\pi} \alpha_k(t) \mathrm{Re}\{f(e^{it})\} dt = c_k, \quad k = 0, 1, \ldots, n$$



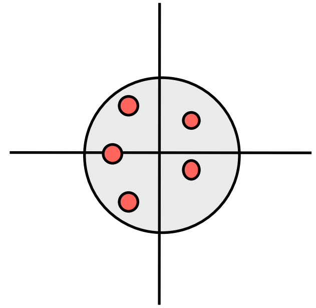
analytic in $\mathbb{D}$

$\mathrm{Re}\{f(z)\} \geq 0$ in $\mathbb{D}$

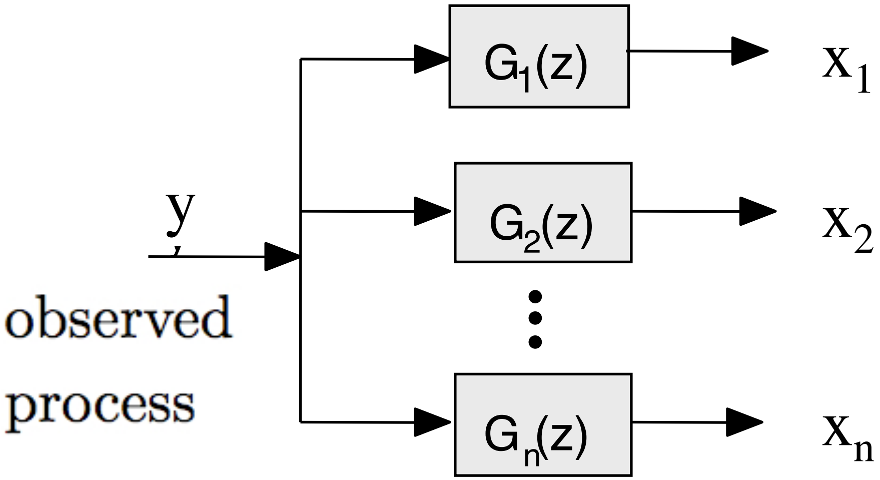
$$\boxed{d\mu = \mathrm{Re}\{f(e^{it}\} dt \in \mathcal{R}_+ \quad \Longleftrightarrow \quad \deg f \leq n}$$

A tunable high resolution spectral estimator (THREE)

Zoom into a selected spectral band by moving interpolation points from the origin closer to the unit circle.



Byrnes-Georgiou-L



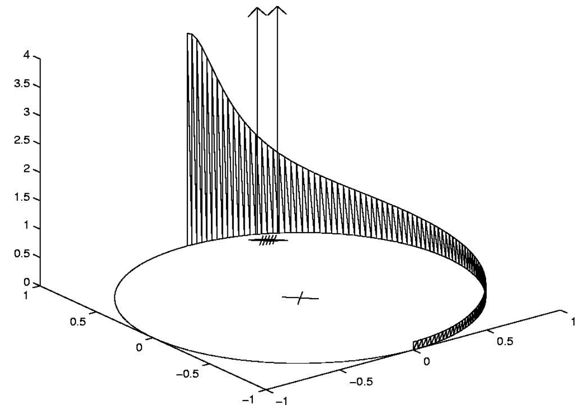
$$G_k(z) = \frac{1}{1 - \bar{z}_k z}$$

$$\boxed{f(z_k) = w_k := \tfrac{1}{2}(1 - z_k^2)E\{x_k^2\}}$$

$$E\{x(t)x(t)^{\mathsf{T}}\} = \left[\frac{w_k + \bar{w}_\ell}{1 - z_k \bar{z}_\ell}\right]_{k,\ell=0}^{n}$$

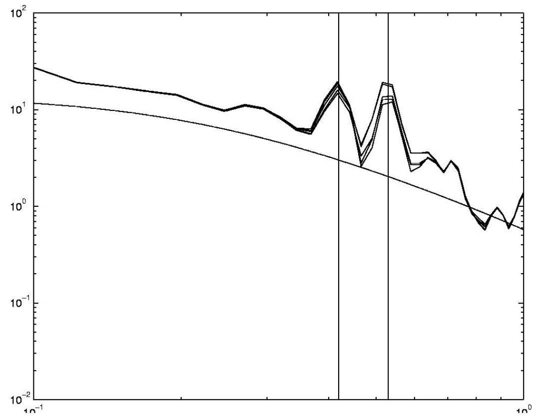
Two sets of tuning parameters:
- filter bank poles
- spectral zeros (P)

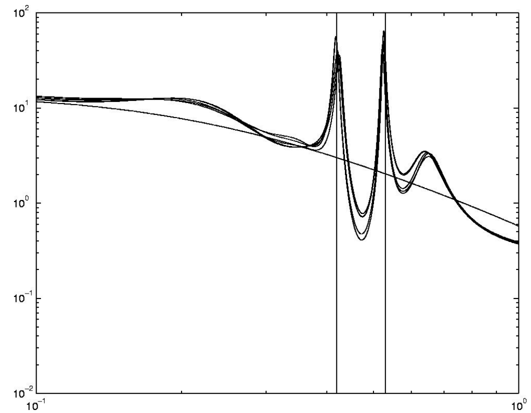
Estimation of spectral lines in colored noise



separation between
spectral lines = 0.11

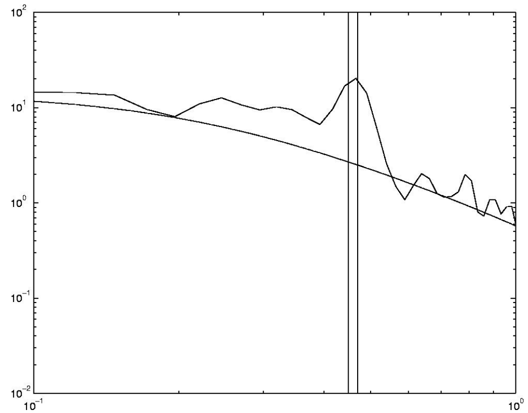
five runs superimposed

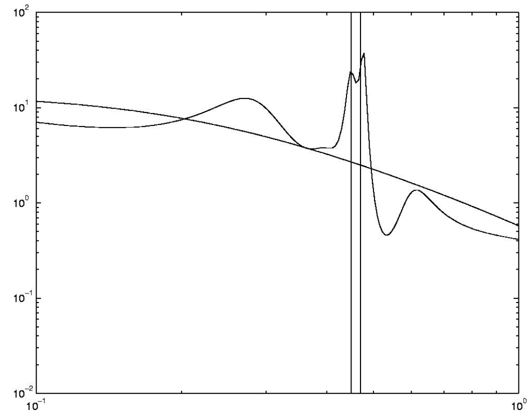




Periodogram (FFT)   THREE (default setting)

separation between
spectral lines = 0.02





Conclusions

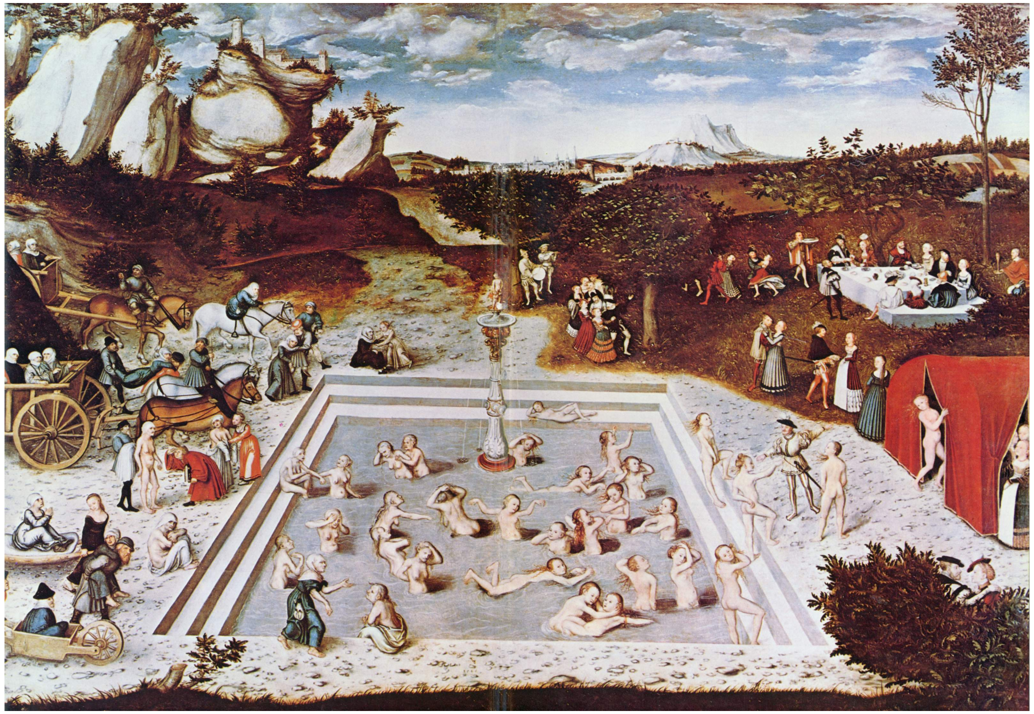
An enhanced theory for generalized moment problems that incorporates rationality constraints prescribed by applications.

- Complete parameterizations of solutions with smooth tuning strategies.
- A global analysis approach that studies the class of solutions as a whole.
- Convex optimization for determining solutions.

Basic observation: Boris is invariant under age

I am looking forward to the 100 year celebration of Boris

I suspect he will look the same then also, because he has found



The Fountain of Youth

Finally a commercial