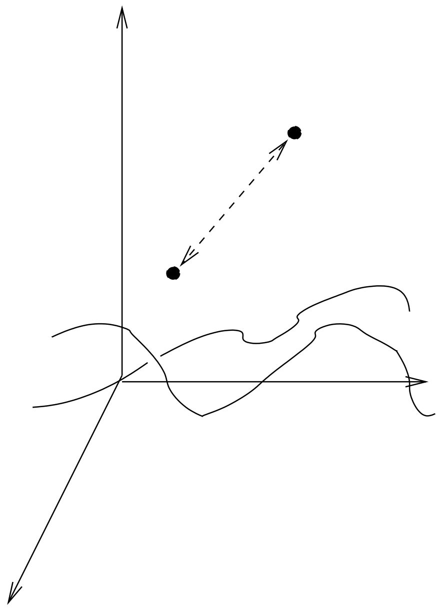
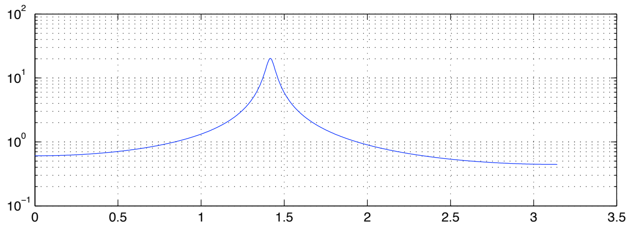
Distances and Riemannian metrics spectral analysis

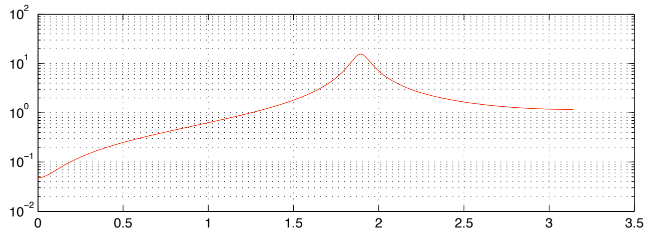
Tryphon Georgiou

Electrical & Computer Engineering
Univ. of Minnesota

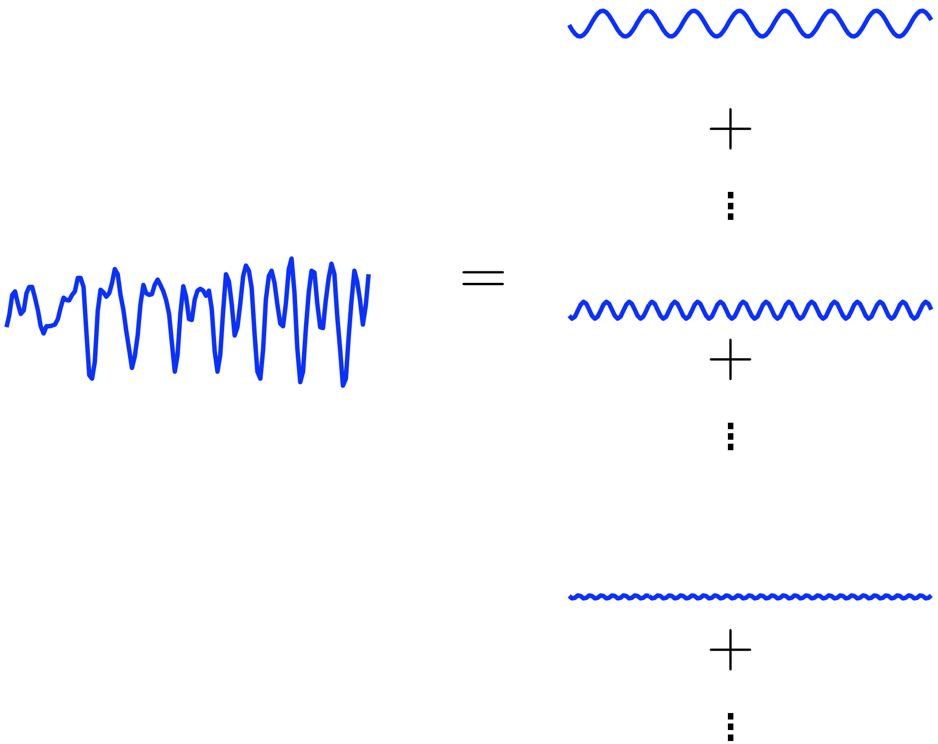
meaning of distances



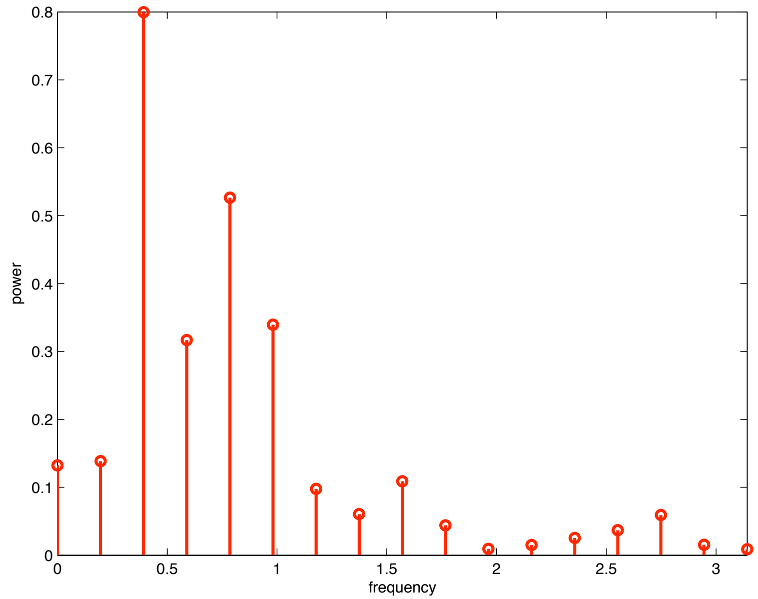




spectral analysis



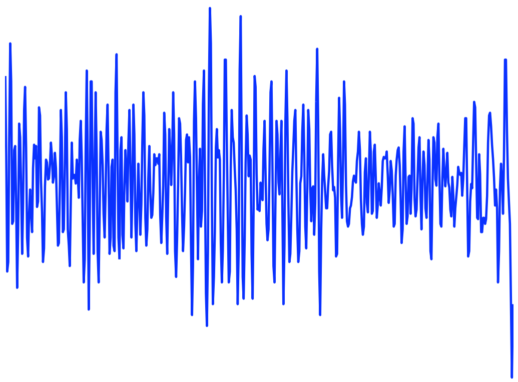
$$\ldots u_{-1}, u_0, u_1, \ldots \qquad u_k = \sum e^{jk\theta_i} X_i$$



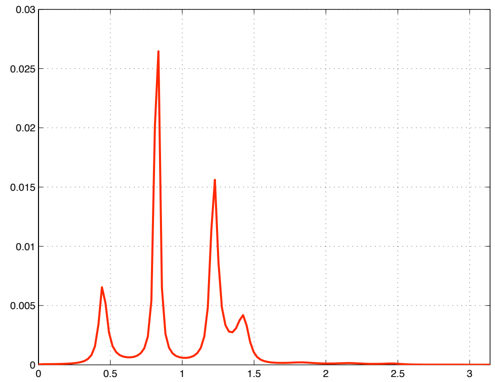
$$u_k = \int e^{jk\theta} \, dX(\theta)$$

$$f(\theta) = E\{|dX(\theta)|^2\}$$

spectral analysis tools



⇒



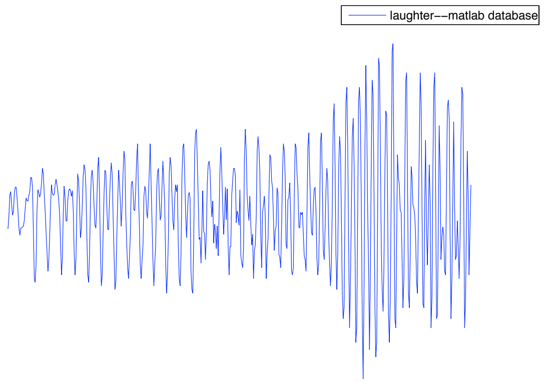
- Fourier transform, periodogram, Blackmann-Tukey

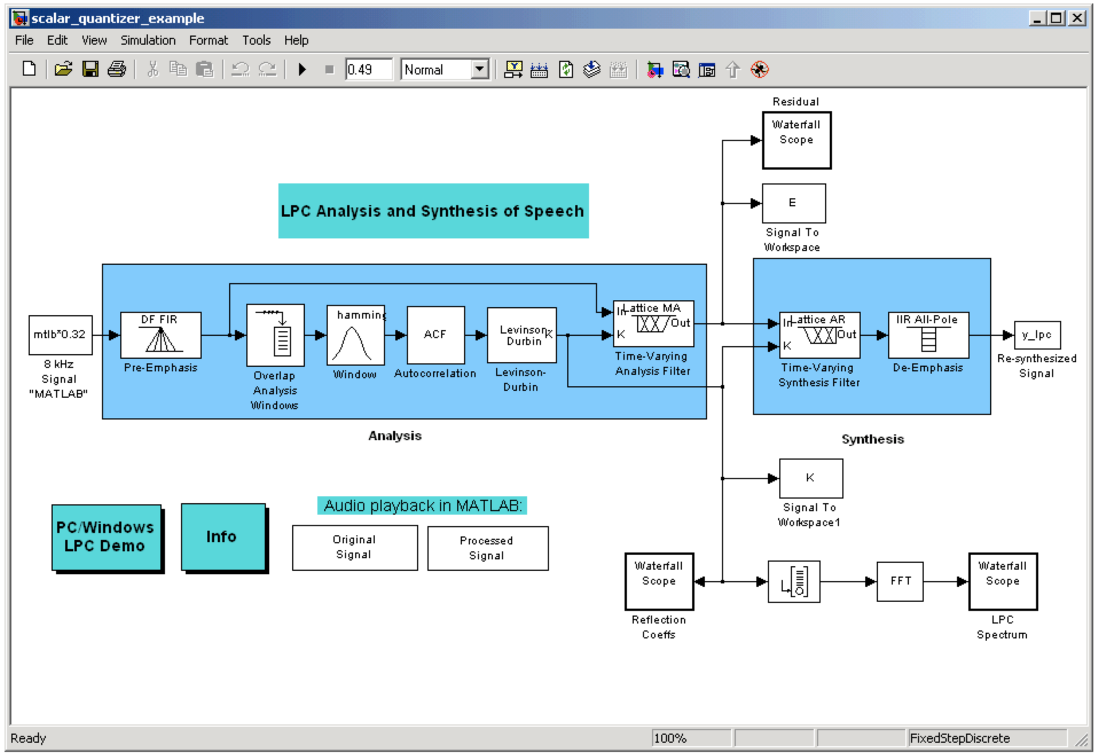
- Levinson, Durbin, Burg, Capon, . . .

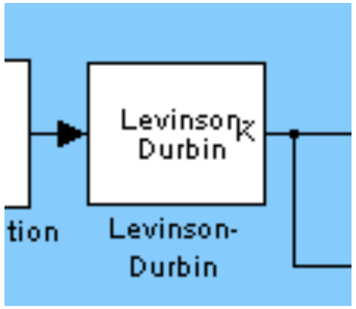
- Subspace methods, Caratheodory, MUSIC, ESPRIT, . . .

a hidden technology (communications)

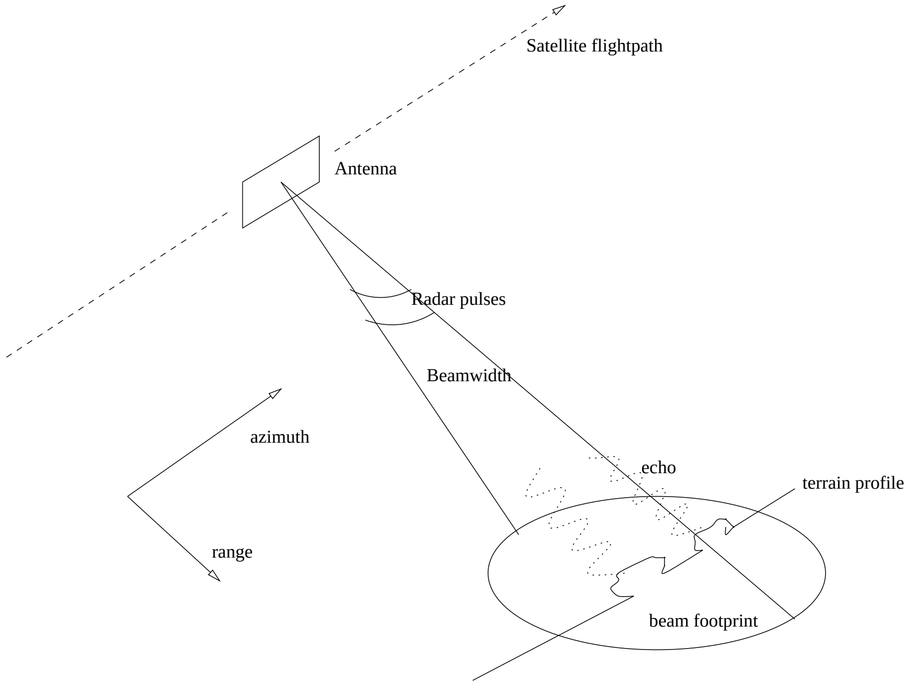
- speech analysis/coding

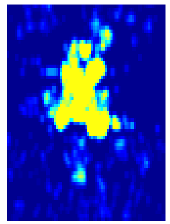






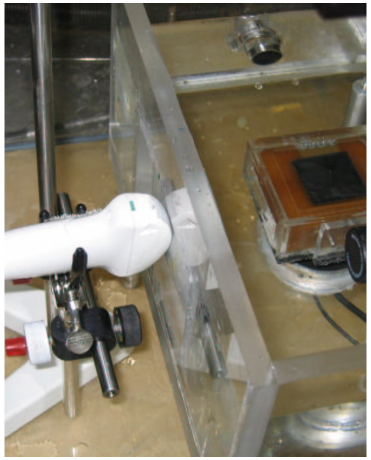
a hidden technology (radar)

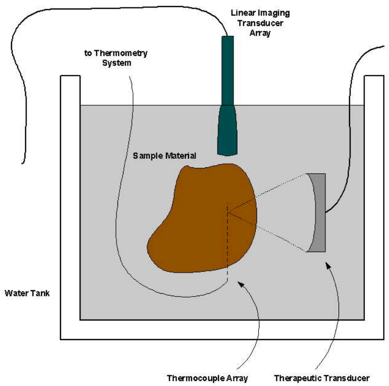




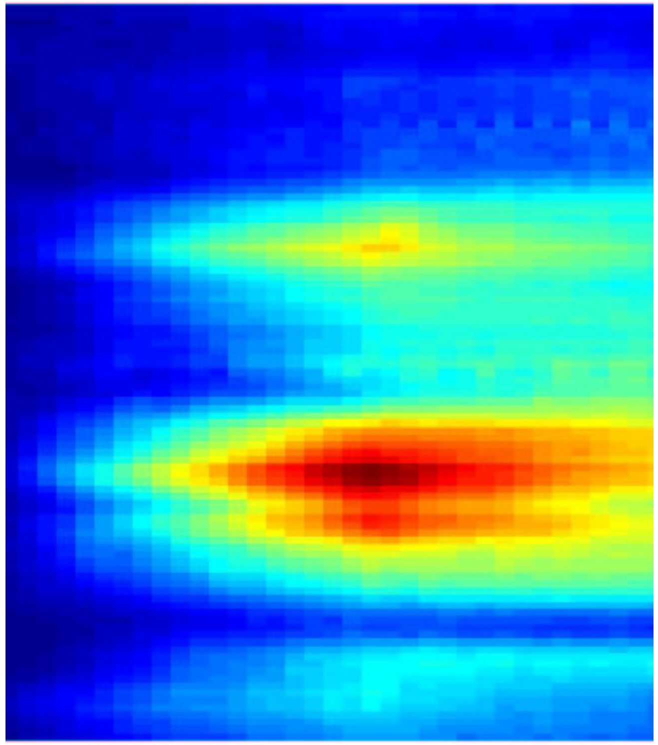
collaboration A. Nasiri-Amini

a hidden technology (medical diagnostics)





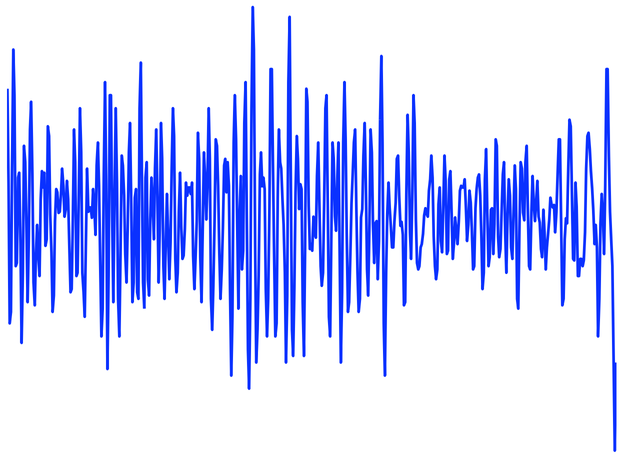
Noninvasive temperature sensing



Temperature field (color coded) position vs. time

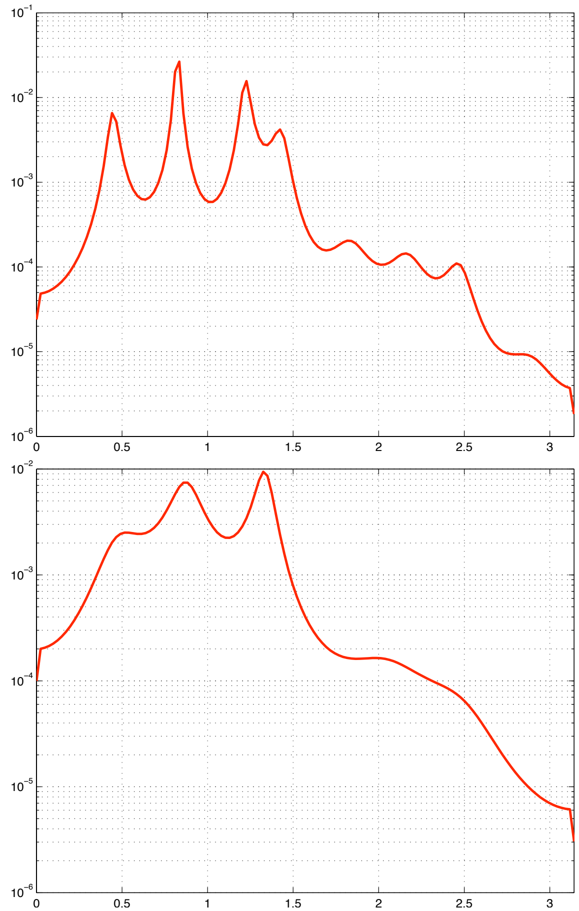
collaboration E. Ebbini
& A. Nasiri Amini

quantitative analysis



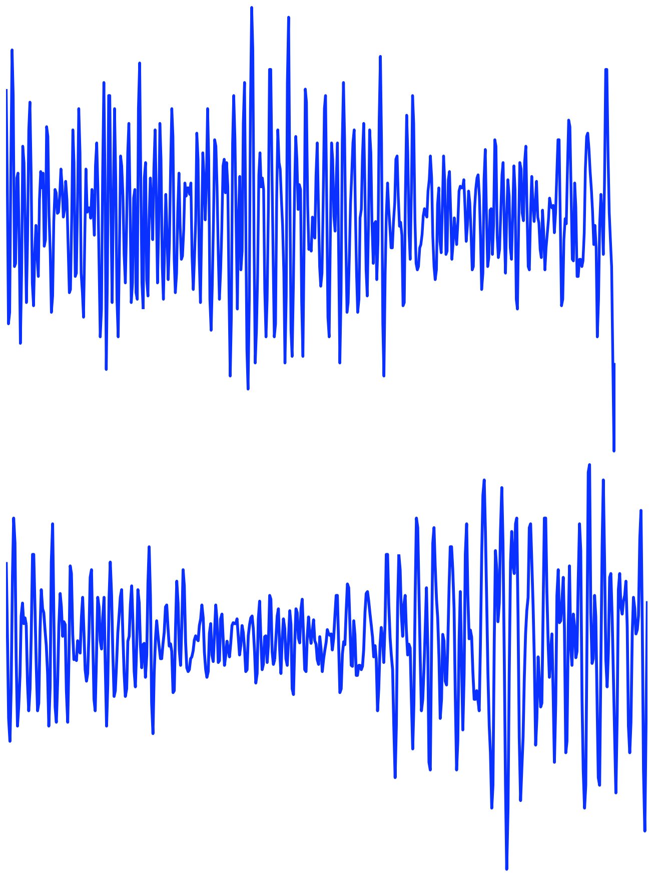
different

methods



How can we compare power spectra?

quantitative analysis

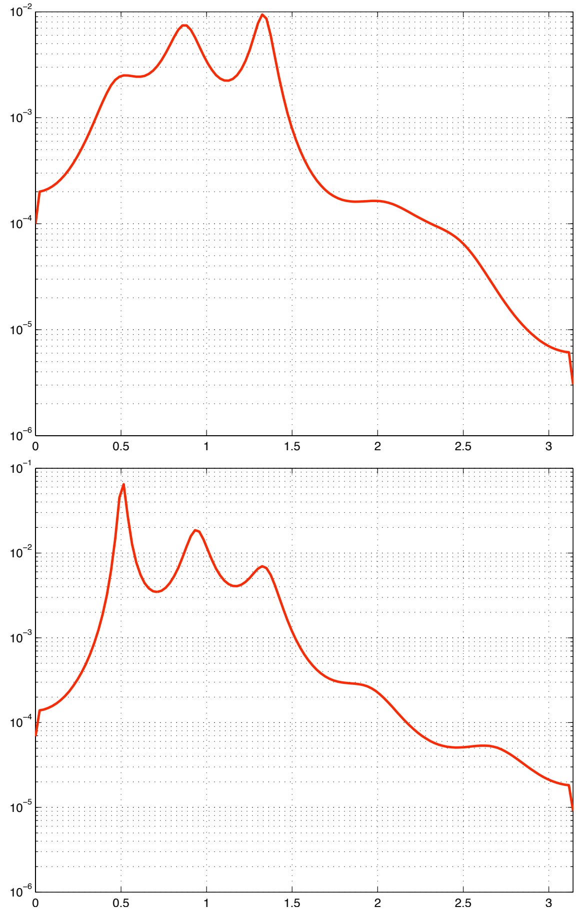


→

same

method

→



How can we compare power spectra?

How can we compare power spectra?

Question:

what is a natural notion of distance between power spectral densities?

Goal

- quantitative spectral analysis:

  compare performance of algorithms

  tune algorithms

  assess uncertainty

  assess affinity between spectra

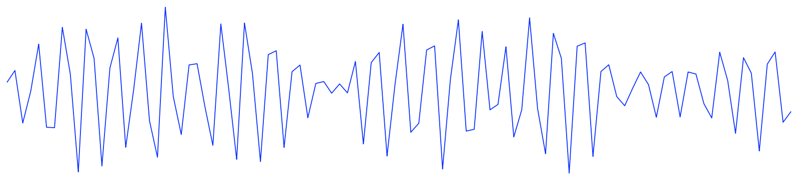
  assess drift, structural changes

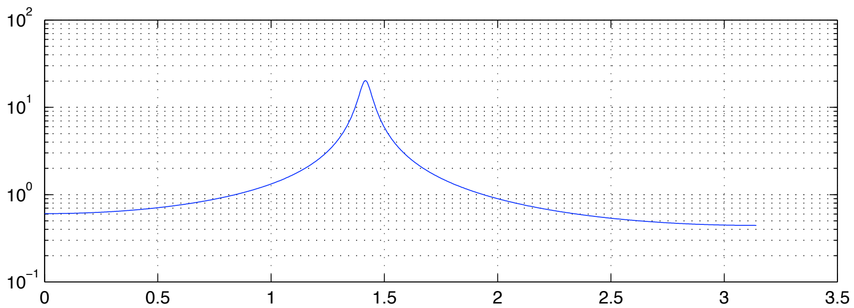
  signal classification in speech analysis, radar, etc. etc.

- system identification

setting

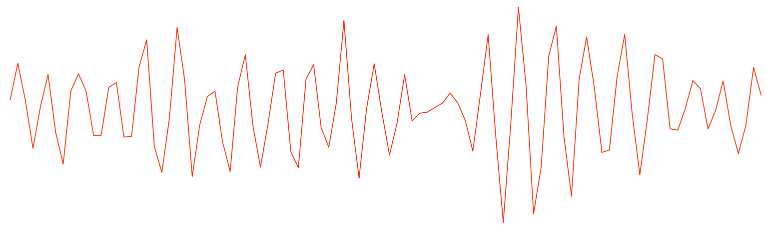
$\ldots u_{-1}, u_0, u_1, u_2, \ldots$

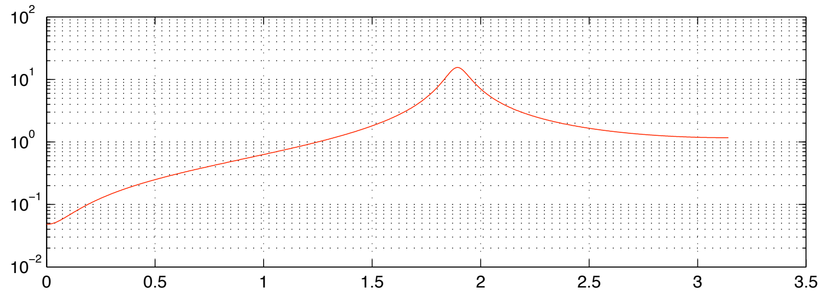




$f_1(\theta),\ \theta \in [0, \pi]$

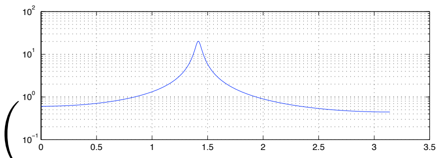
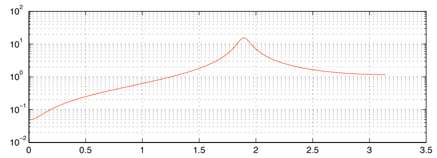
$\ldots u_{-1}, u_0, u_1, u_2, \ldots$





$f_2(\theta),\ \theta \in [0, \pi]$

what is it we would like to have?

distance( , )

$f_1(\theta)$, $f_2(\theta)$

- metric

- meaningful & natural

candidate distances?

$$\|f_1 - f_2\|^2_{\mathrm{rms}} = \int |f_1(\theta) - f_2(\theta)|^2 d\theta$$

no interpretation. . .

$f_1 - f_2$ is not a “signal”

$\int f_1^2$ is not “energy”

candidate distances?

Itakura-Saito, Ali-Silvey, etc.

F strictly convex $\stackrel{\text{Bregman}}{\Longrightarrow} d(f_1, f_2) = \int [F(f_1) - F(f_2) - F'(f_1 - f_2)] \, d\theta$

a possibility, but no interpretation. . .

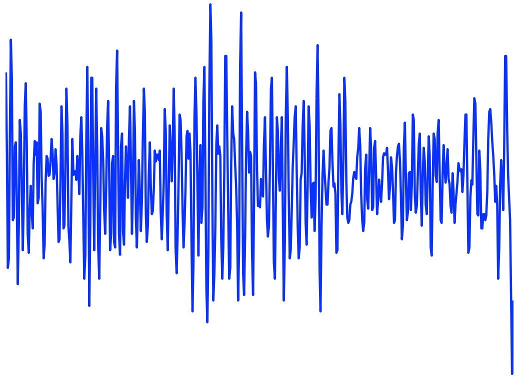
Kullback-Leibler:

$\int f_1 \left(\log(f_1) - \log(f_2) \right) d\theta$

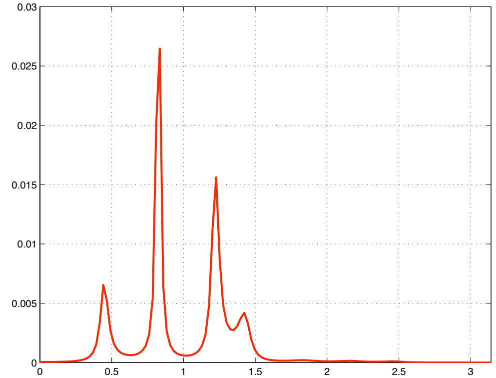
$\| \log(f_1) - \log(f_2) \|_{\text{rms}}$

$\int | \log(f_1) - \log(f_2) | d\theta, \ldots$

notation



$\Rightarrow$



$$\{\ldots, u_0, u_1, \ldots\} \quad \Rightarrow \quad R_k := E\{u_\ell u_{\ell-k}\} \quad \Rightarrow \quad R_k = \int e^{jk\theta} f(\theta) d\theta$$

statistics

. . . whenever necessary: $\{u_{f_1,k}\}$, and $\{u_{f_2,k}\}$

Least-variance approximation

One-step-ahead prediction: $u_0 - u_{0|\text{past}}$

with $u_{0|\text{past}} := \sum_{k>0} \alpha_k u_{-k}$

$$E\{|u_0 - u_{0|\text{past}}|^2\} = \text{variance of prediction error}$$

Szegö-Kolmogorov theorem

$$\inf_{\alpha} E\{|u_0 - \sum_{k>0} \alpha_k u_{-k}|^2\} = \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \log f(\theta) d\theta\right\}$$

when $\log f \in L_1$, and zero otherwise.

geometric mean

Discrete:

$$\exp\{\frac{1}{3}(\log f_1 + \log f_2 + \log f_3)\} = \sqrt[3]{f_1 f_2 f_3}$$

Continuous analog:

$$\exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \log f(\theta) d\theta\right\}$$

the optimal predictor

$$f(\theta) = \frac{g_f}{|a_f(e^{j\theta})|^2}$$

$a_f(z)$ invertible & normalized so that $a_f(0) = 1$

$$g_f = \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\log(f(\theta))d\theta\right)$$

$$u_0 - \hat{u}_{0|\text{past}} \mapsto a_f(z) = 1 - a_{f,1}z - a_{f,2}z^2 + \ldots$$

Degradation of prediction error variance

using f_2 to design a predictor
and then comparing how it performs on f_1

$$\rho(f_1, f_2) := \frac{E\{|u_{f_1,0} - \sum_{\ell=1}^{\infty} a_{f_2,\ell}\, u_{f_1,-\ell}|^2\}}{E\{|u_{f_1,0} - \sum_{\ell=1}^{\infty} a_{f_1,\ell}\, u_{f_1,-\ell}|^2\}}$$

substituting. . .

$$\begin{aligned}\rho(f_1, f_2) &= \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |a_{f_2}(e^{j\theta})|^2 f_1(\theta) d\theta\right) / g_{f_1} \\ &= \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{f_1(\theta)}{f_2(\theta)} d\theta\right) \frac{g_{f_2}}{g_{f_1}}.\end{aligned}$$

Degradation. . .

$$\rho(f_1, f_2) = \left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{f_1(\theta)}{f_2(\theta)}d\theta\right)\frac{1}{\exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\log(\frac{f_1(\theta)}{f_2(\theta)})d\theta\right)}.$$

arithmetic over *geometric* mean (≥ 1)

$$\delta_{\mathrm{a/g}}(f_1, f_2) := \log\rho(f_1, f_2)$$

$$= \log\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{f_1(\theta)}{f_2(\theta)}d\theta\right) - \frac{1}{2\pi}\int_{-\pi}^{\pi}\log(\frac{f_1(\theta)}{f_2(\theta)})d\theta \quad (\geq 0)$$

- *not a metric*

Symmetrization...

$$\delta(f_1, f_2) := \delta_{\mathrm{a/g}}(f_1, f_2) + \delta_{\mathrm{a/g}}(f_2, f_1)$$

$$= \log\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{f_1(\theta)}{f_2(\theta)}d\theta\right) + \log\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{f_2(\theta)}{f_1(\theta)}d\theta\right)$$

$$= \log\left(\frac{\text{arithmetic mean of } \frac{f_1(\theta)}{f_2(\theta)}}{\text{harmonic mean of } \frac{f_1(\theta)}{f_2(\theta)}}\right)$$

- $\delta(f_1, f_2) \geq 0$ unless $\frac{f_1(\theta)}{f_2(\theta)} =$ constant
- still not a metric

Riemannian metric

$$\delta\,(f, f+\Delta) = o(\|\Delta\|^2)$$

$\Rightarrow$

$$g_f(\Delta) := \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{\Delta(\theta)}{f(\theta)}\right)^2\frac{d\theta}{2\pi} - \left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\Delta(\theta)}{f(\theta)}\frac{d\theta}{2\pi}\right)^2$$

mean-square vs. arithmetic-mean squared

Geodesics

seeking a path f_τ ($\tau \in [0,1]$) between f_0, f_1 of minimal length

But for us, each point represents a different power spectral density

Geodesics

seeking a path f_τ ($\tau \in [0,1]$) between f_0, f_1 of minimal length

$$\sqrt{2}\int_0^1 \sqrt{\delta\left(f_\tau, f_{\tau+d\tau}\right)} = \int_0^1 \sqrt{g_{f_\tau}(\frac{\partial f_\tau}{\partial \tau})}d\tau$$

Euler-Lagrange:

$$\frac{\partial L}{\partial x_\tau} - \frac{\partial}{\partial \tau}\frac{\partial L}{\partial \dot{x}_\tau} = 0$$

$$L(x_\tau, \dot{x}_\tau, \tau) := \sqrt{\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\dot{x}_\tau(\theta)\right)^2 d\theta - \left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\dot{x}_\tau(\theta)d\theta\right)^2}$$

$x_\tau = \log(f_\tau)$ and $\dot{x}_\tau := \partial x_\tau / \partial \tau$

characterization of geodesics

the geodesics are exponential families:

$$f_\tau(\theta) = f_0(\theta)\left(\frac{f_1(\theta)}{f_0(\theta)}\right)^\tau, \ \tau \in [0,1]$$

or, logarithmic intervals:

$$f_\tau(\theta) = e^{(1-\tau)\log(f_0(\theta))+\tau\log(f_1(\theta))}, \ \tau \in [0,1]$$

Geodesic distance

the path-length along the logarithmic interval connecting f_0 and f_1 is

$$d_g(f_0, f_1) := \sqrt{\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\log\frac{f_1(\theta)}{f_0(\theta)}\right)^2 d\theta - \left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\log\frac{f_1(\theta)}{f_0(\theta)}d\theta\right)^2}$$

mean square of log vs. arithmetic mean of log (“variance”)

(pseudo) metric properties

$d_g(f_0, f_1) \geq 0$ ✓

$d_g(f_0, f_1) = d_g(f_1, f_0)$ ✓

$d_g(f_1, f_2) + d_g(f_2, f_3) \geq d_g(f_1, f_3)$ ✓

using $\log \frac{f_1}{f_3} = \log \frac{f_1}{f_2} + \log \frac{f_2}{f_3}$, and $\alpha := \log(f_1/f_2)$, $\beta := \log(f_2/f_3)$
rewrite $d_g(f_1, f_2) + d_g(f_2, f_3) \geq d_g(f_1, f_3)$:

$$\sqrt{\frac{1}{2\pi}\int_{-\pi}^{\pi} \alpha^2 \frac{d\theta}{2\pi} - (\frac{1}{2\pi}\int_{-\pi}^{\pi} \alpha \frac{d\theta}{2\pi})^2}\sqrt{\frac{1}{2\pi}\int_{-\pi}^{\pi} \beta^2 \frac{d\theta}{2\pi} - (\frac{1}{2\pi}\int_{-\pi}^{\pi} \beta \frac{d\theta}{2\pi})^2}$$
$$\geq \frac{1}{2\pi}\int_{-\pi}^{\pi} (\alpha\beta) \frac{d\theta}{2\pi} - \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \alpha \frac{d\theta}{2\pi}\right)\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \beta \frac{d\theta}{2\pi}\right)$$

...

Information geometry – parallels

degradation of performance – Kullback-Leibler distance (not a metric)

$$\left(- \int p_1 \log(p_0) \right) - \left(- \int p_1 \log(p_1) \right)$$

Riemannian metric – Fisher information metric

$$\begin{aligned} g_{\text{Fisher},p}(\Delta) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{\Delta(\theta)}{p(\theta)} \right)^2 p(\theta) \frac{d\theta}{2\pi} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\Delta(\theta)^2}{p(\theta)} \frac{d\theta}{2\pi} \end{aligned}$$

Information geometry – parallels (cont.)

geodesics, geodesic distance

$$p \mapsto \sqrt{p} \in \mathrm{Sphere}$$

geodesics: great circles

geodesic distance is the arclength: relates to Battacharrya distance. . .

back to “spectral geometry” – alternatives

Smoothing: $u_0 - \hat{u}_{0|\text{past \& future}}$

with $\hat{u}_{0|\text{past \& future}} := \sum_{k \neq 0} \beta_k u_{-k}$

$$E\{|u_0 - \hat{u}_{0|\text{past \& future}}|^2\} = \|1 - \sum_{k \neq 0} \beta_k e^{jk\theta}\|_{d\mu}^2$$

Least-variance smoothing

$$\inf_{\beta} \|1 - \sum_{k \neq 0} \beta_k e^{jk\theta}\|^2_{d\mu} = \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta)^{-1} d\theta \right)^{-1} =: h_f$$

when $f^{-1} \in L_1$, and zero otherwise.

http://arxiv.org/abs/math/0601648

harmonic mean

Discrete:

$$\frac{1}{\frac{1}{3}(\frac{1}{f_1} + \frac{1}{f_2} + \frac{1}{f_3})}$$

Continuous analog:

$$\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta)^{-1} d\theta \right)^{-1}$$

harmonic mean $\leq$ geometric mean

$$\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)^{-1} d\theta\right)^{-1} \leq \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left(f(\theta)\right) d\theta\right).$$

Degradation of "smoothing" variance

$$\rho_{\text{smooth}}(f_1, f_2) := \frac{E\{|u_{f_1,0} - \sum_{\ell \neq 0} b_{f_2,\ell}\, u_{f_1,-\ell}|^2\}}{E\{|u_{f_1,0} - \sum_{\ell \neq 0} b_{f_1,\ell}\, u_{f_1,-\ell}|^2\}}$$

$$\rho_{\text{smooth}}(f_1, f_2) = \left(\frac{\sqrt{\frac{1}{2\pi}\int_{-\pi}^{\pi} \left(\frac{f_1(\theta)}{f_2(\theta)}\right)^2 d\phi_1(\theta)}}{\frac{1}{2\pi}\int_{-\pi}^{\pi} \left(\frac{f_1(\theta)}{f_2(\theta)}\right) d\phi_1(\theta)} \right)^2$$

where

$$d\phi_1(\theta) := \frac{f_1(\theta)^{-1} d\theta}{\frac{1}{2\pi}\int_{-\pi}^{\pi} f_1(\theta)^{-1} d\theta}$$

Then:

$$\delta_{\text{smooth}}(f_1, f_2) \; = \; \log(\rho_{\text{smooth}}(f_1, f_2))$$

$$= \log \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{f_1(\theta)}{f_2(\theta)} \right)^2 d\phi_1(\theta) \right) - \log \left(\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{f_1(\theta)}{f_2(\theta)} \right) d\phi_1(\theta) \right)^2 \right)$$

weighted mean-square vs. $(\text{weigted mean})^2$

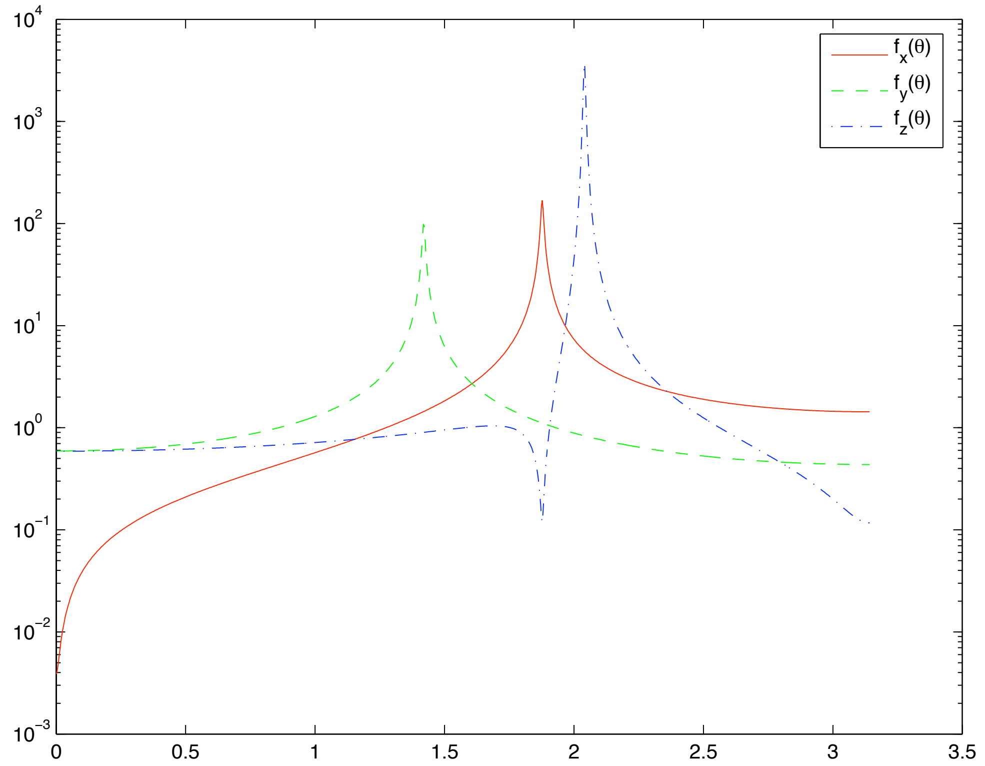
metric, geodesics etc. . . .

Example: f_x (–), f_y (- -), f_z (-.-)

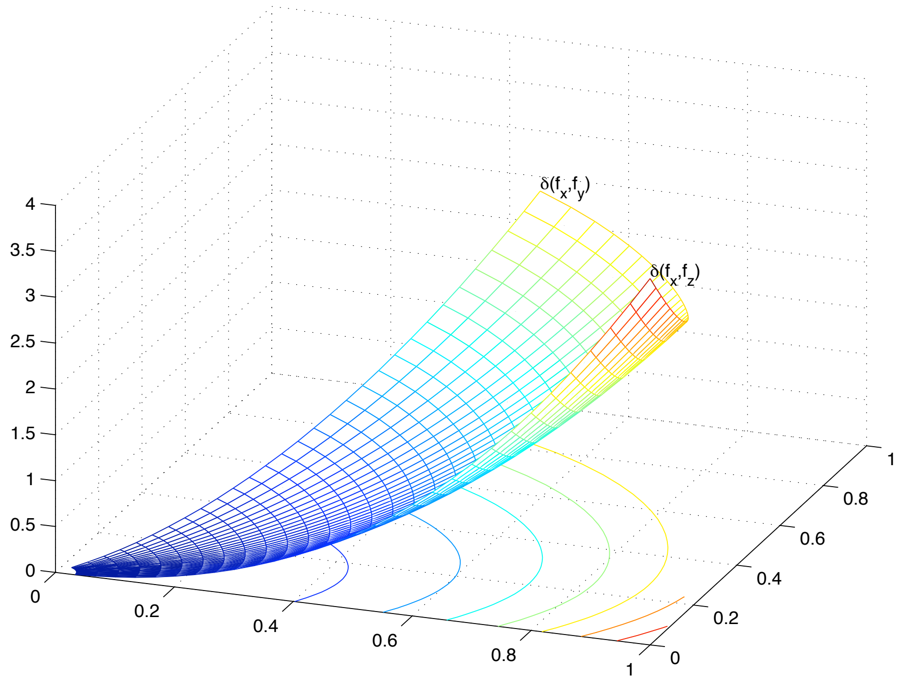
$$f_x(\theta) = \left| \frac{(z - .99)}{(z^2 + .6z + .99)} \right|^2_{z=e^{j\theta}}$$

$$f_y(\theta) = \left| \frac{1}{(z^2 - .3z + .99)} \right|^2_{z=e^{j\theta}}$$

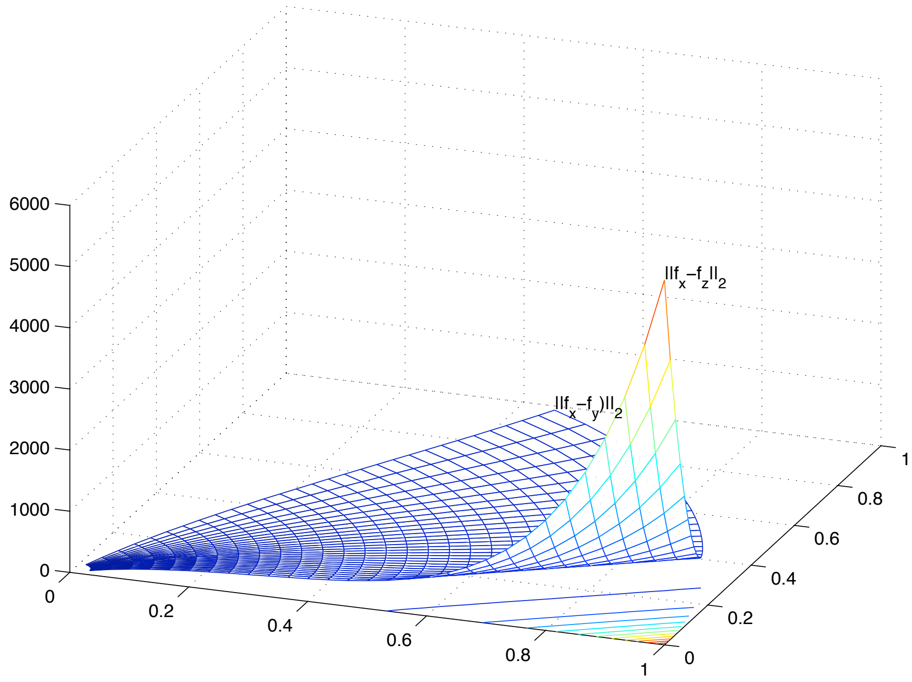
$$f_z(\theta) = \left| \frac{(z + .9)(z^2 + .6z + .99)}{(z^2 + .9z + .99)(z^2 + .9z + .99)} \right|^2_{z=e^{j\theta}}$$



Example: distances to vertex f_x

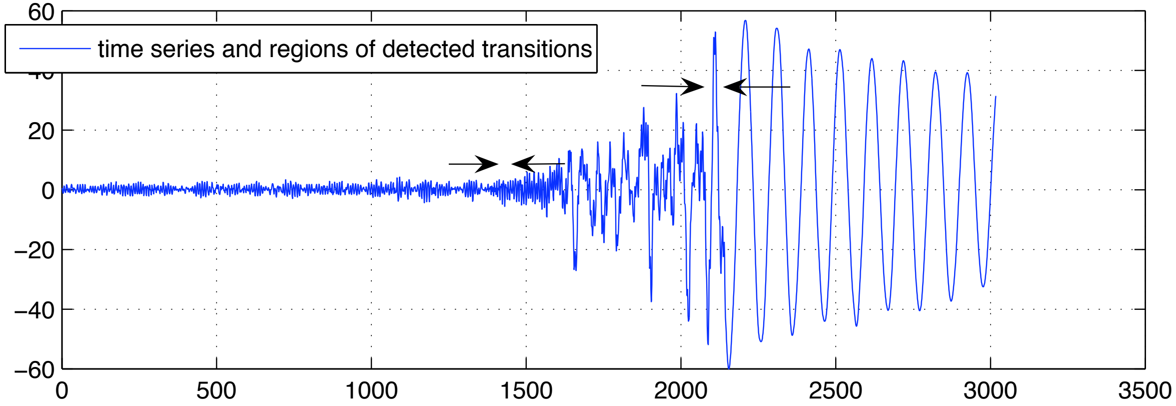


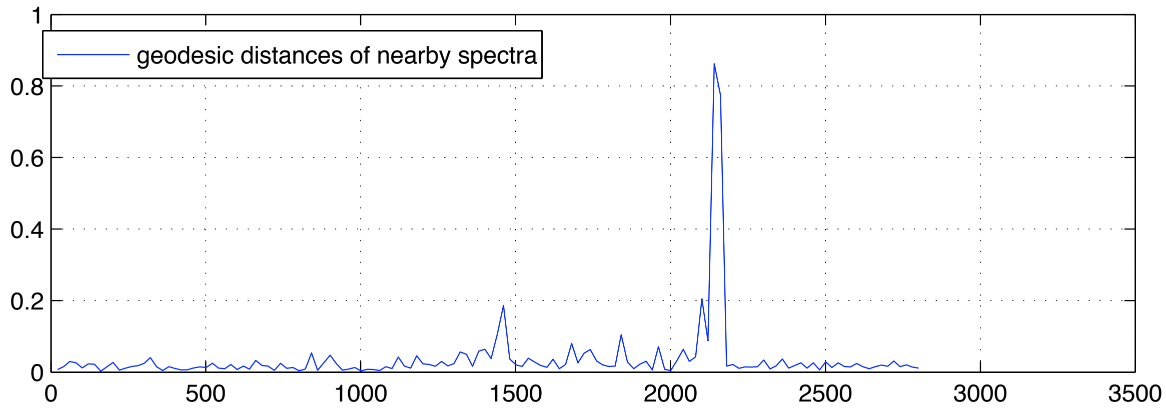
distances in $\delta(\cdot,\cdot)$



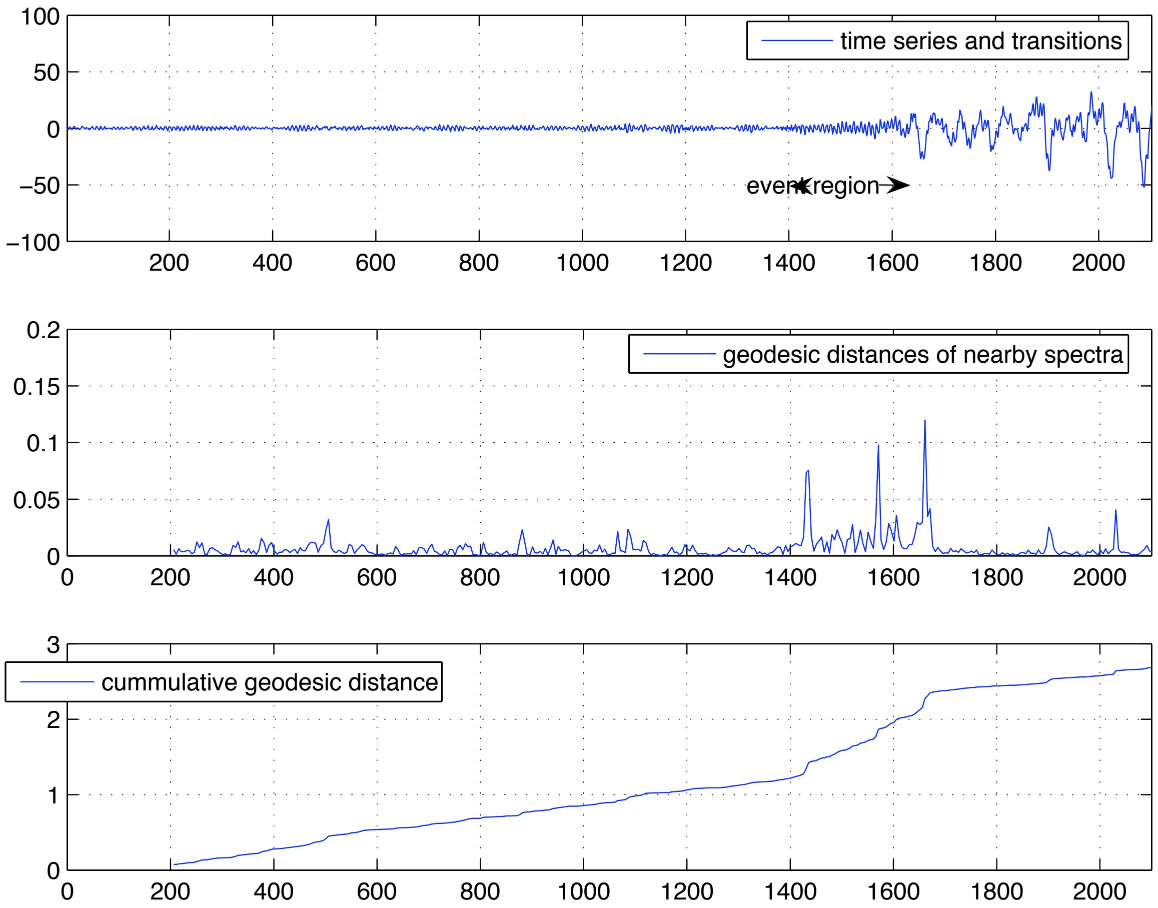
distances in $\|\cdot\|_{\mathrm{rms}}$

Applications – detect drift/transitions





Applications – detect drift/transitions



Concluding thoughts

meaning of distances
in spectral analysis

Degradation of performance

(i) Riemannian metric
(ii) explicit geodesics
(iii) a geodesic distance (metric)

tools for quantitative analysis:
performance of algorithms
signal classification, etc.

metrics in the form of
generalized means