The meaning of Distances in Spectral Analysis

46$^{\mathrm{th}}$ IEEE Conference on Decision & Control

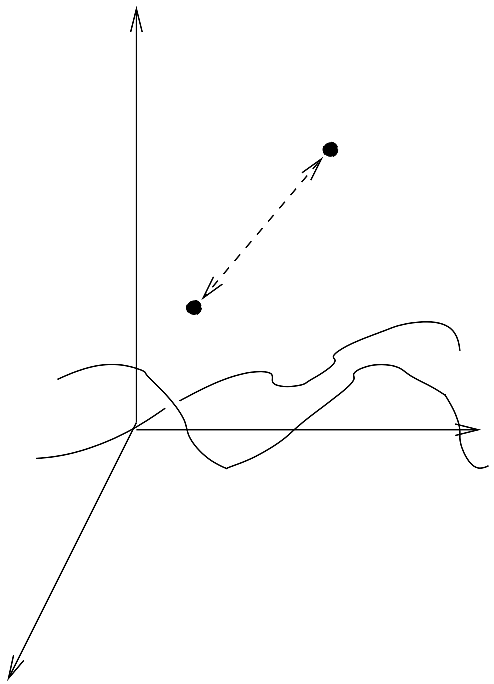
Plenary presentation

Tryphon Georgiou

Electrical & Computer Engineering

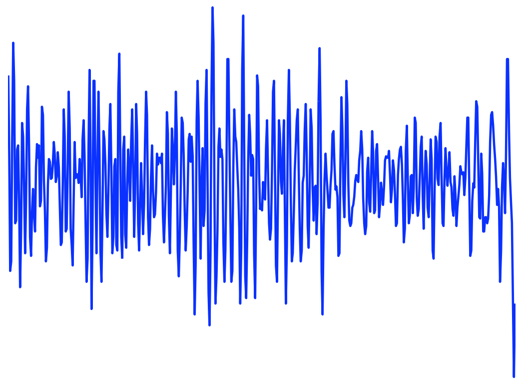
University of Minnesota

Meaning of distances

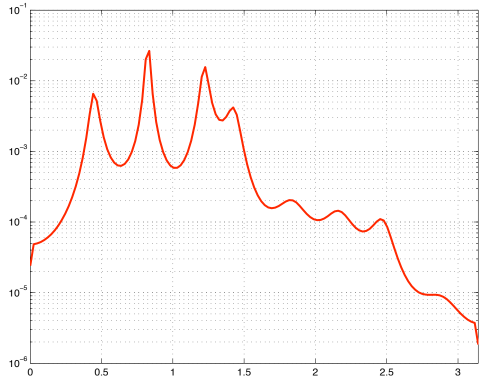


- maximal separation (L_∞)
- energy-like content (L_2)
- integral of flow-rate (L_1)

Power spectra

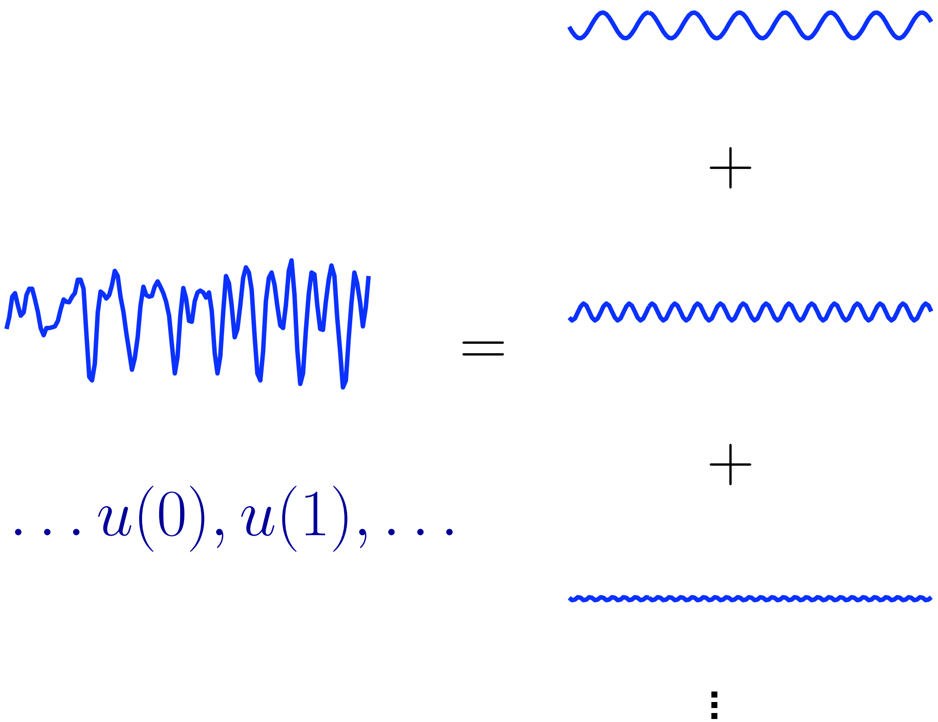


$\Longrightarrow$

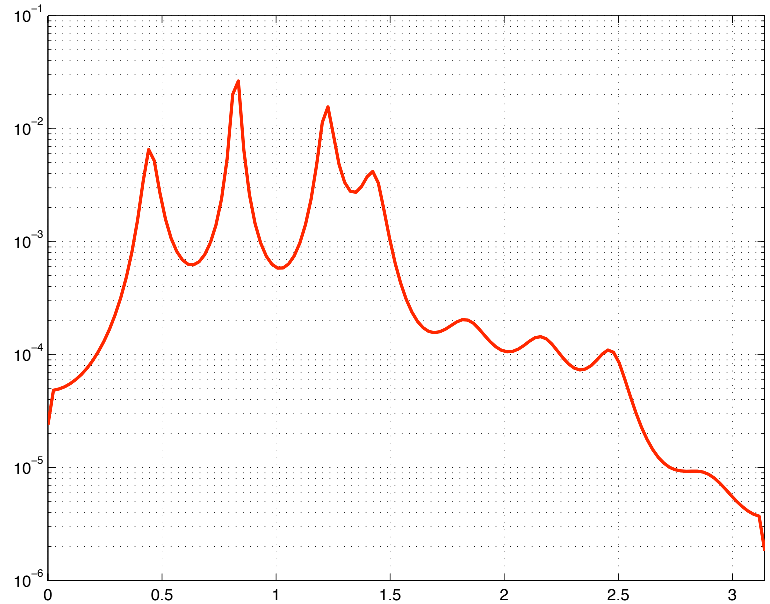


Periodogram, Blackman-Tukey, Levinson, Durbin, Burg, . . .

Spectral analysis



$\ldots u(0), u(1), \ldots$

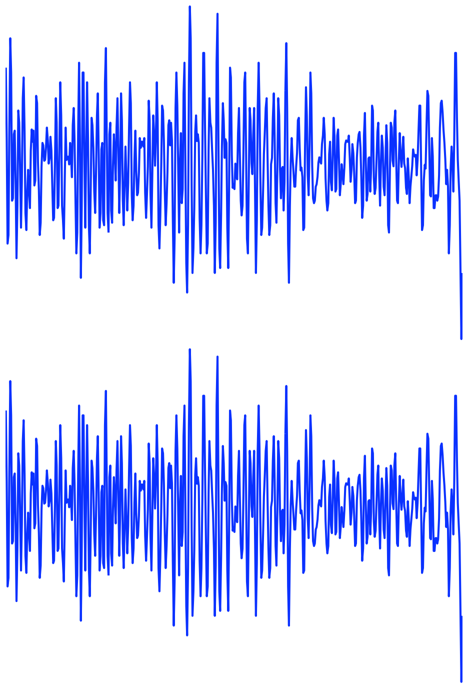


$$u(k) = \int e^{jk\theta} dX(\theta)$$

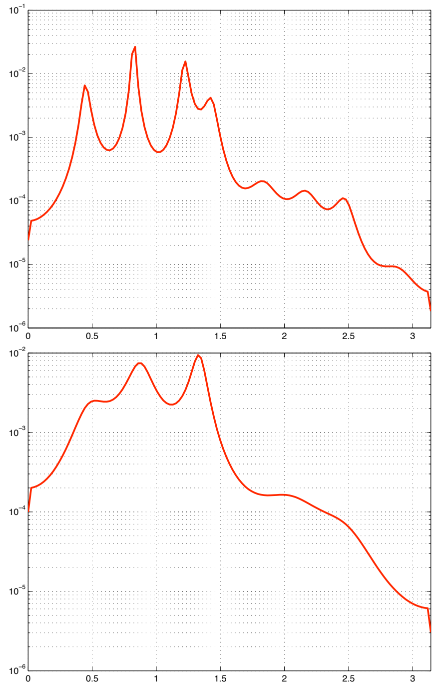
$$E\{u(k)u(k+\ell)\} = \int e^{j\ell\theta} f(\theta) d\theta$$

Signals vs. power densities

time-signals



power distributions

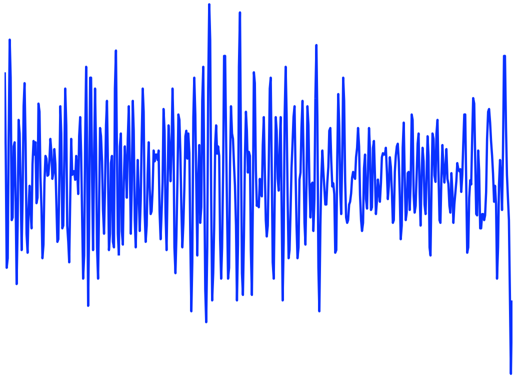


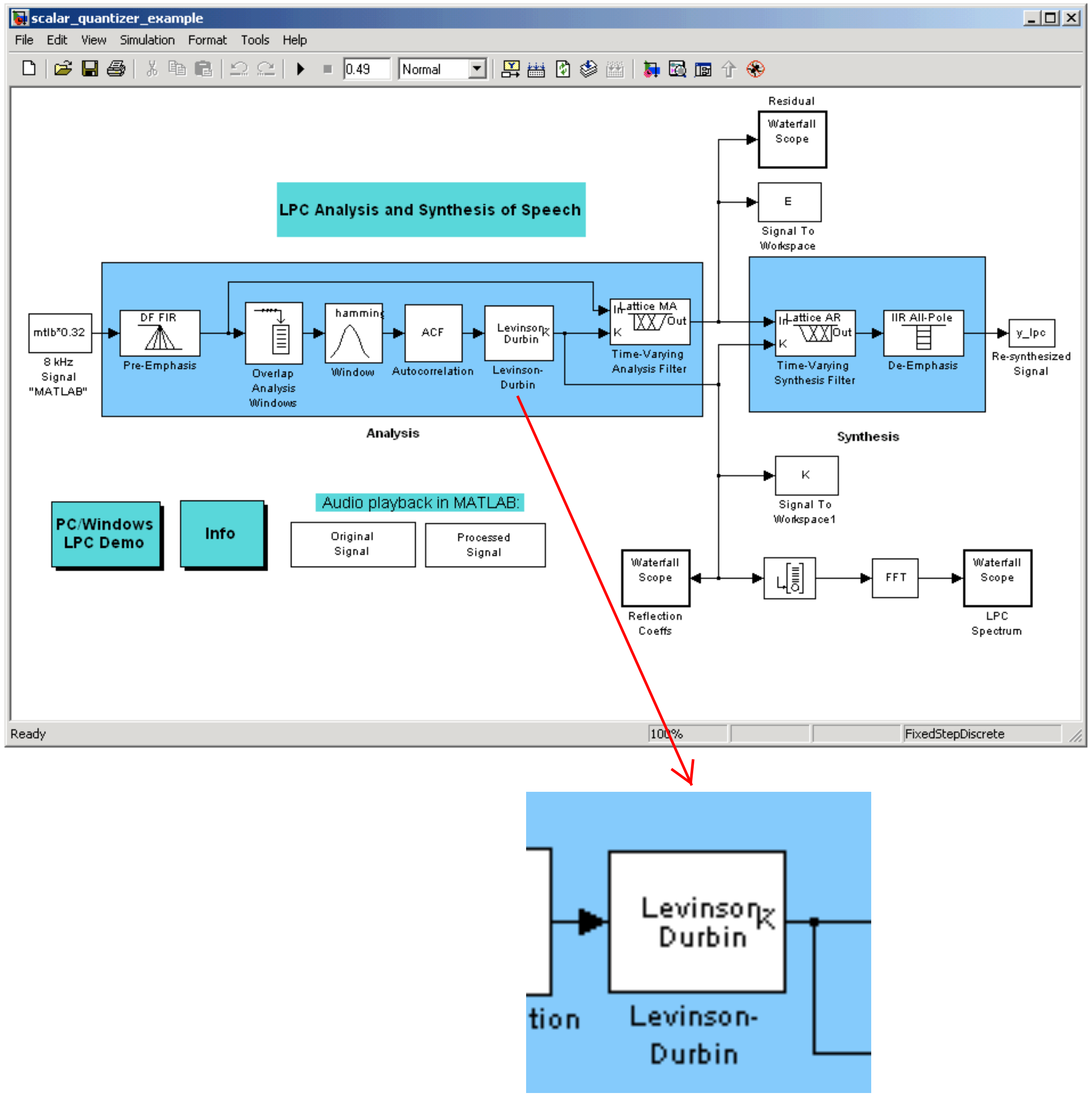
$(u_1 - u_2)$ “error signal”

$(f_1 - f_2)$ is not a “signal”

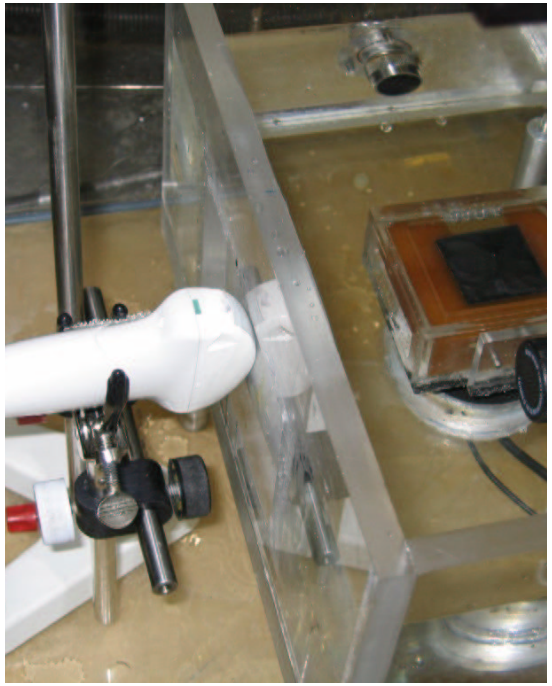
Communications

Speech analysis/coding

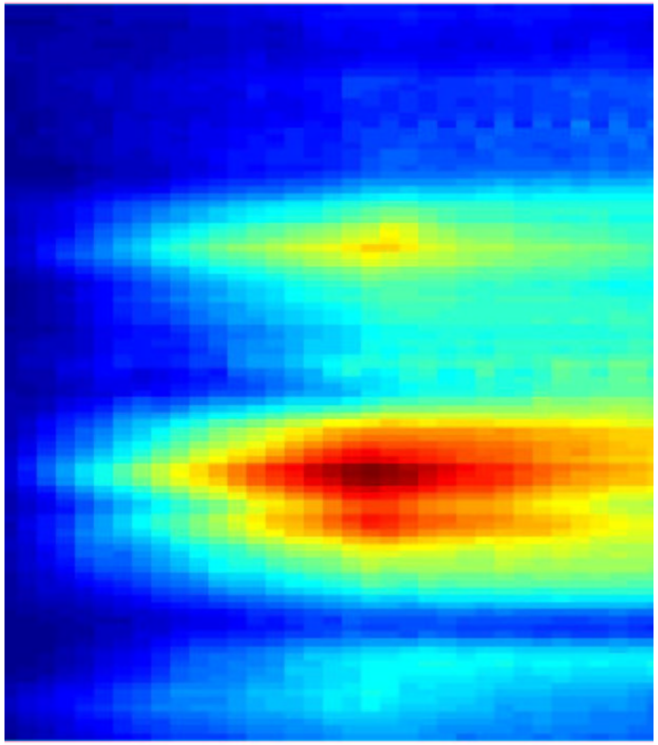




Medical diagnostics



Noninvasive temperature sensing

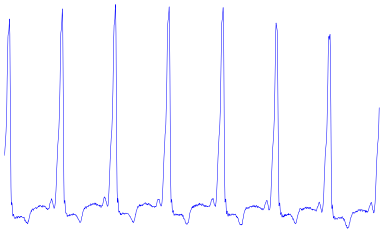


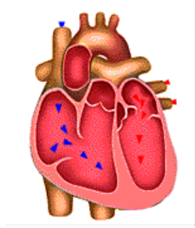
Temperature field

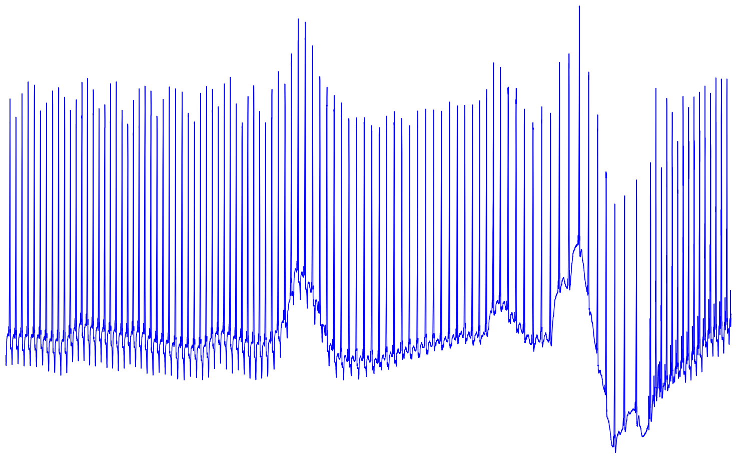
with E. Ebbini & A.N. Amini

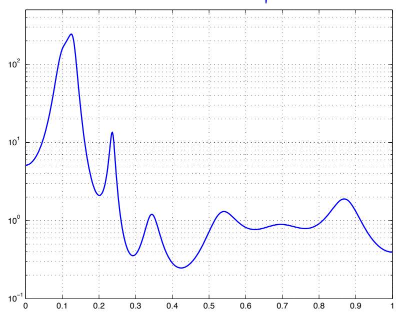
In IEEE Trans. on Biomedical Engineering, 2005

Medical diagnostics

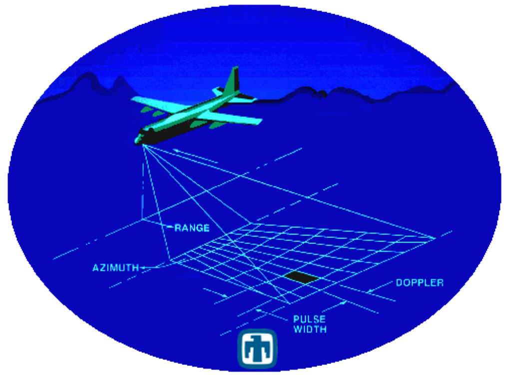




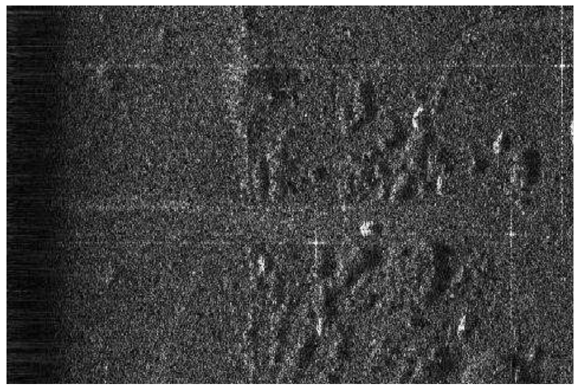




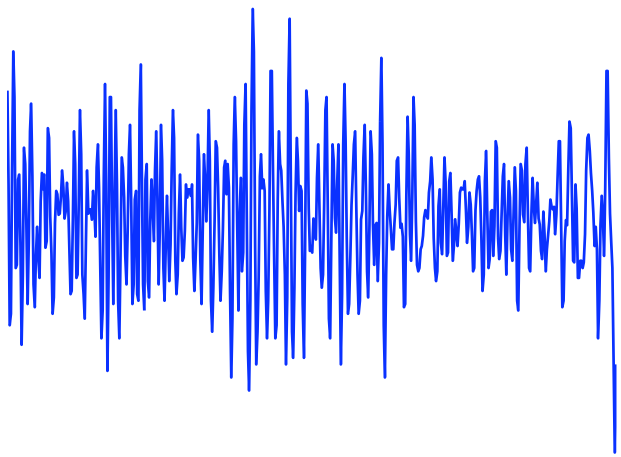
Radar (SAR)



http://www.sandia.gov/radar/images/3dsar.gif

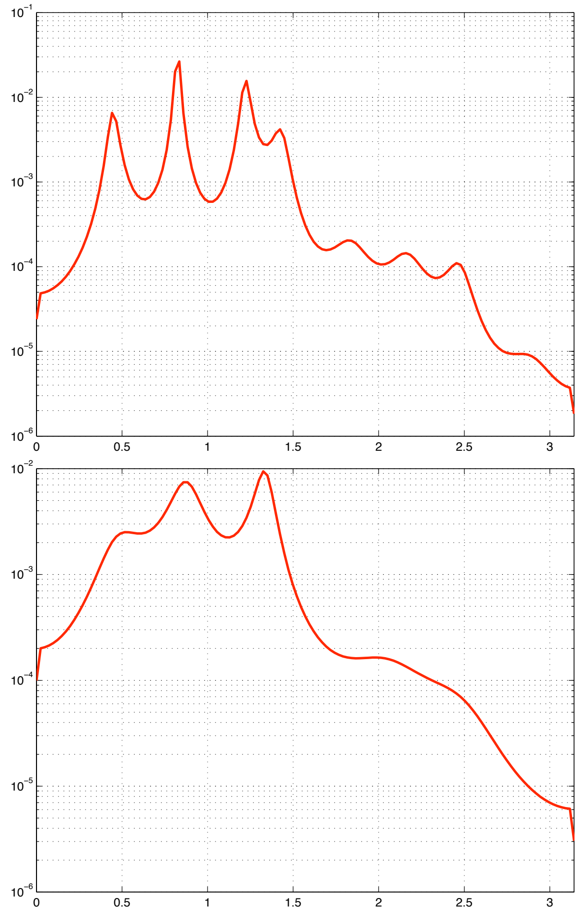


Quantitative analysis



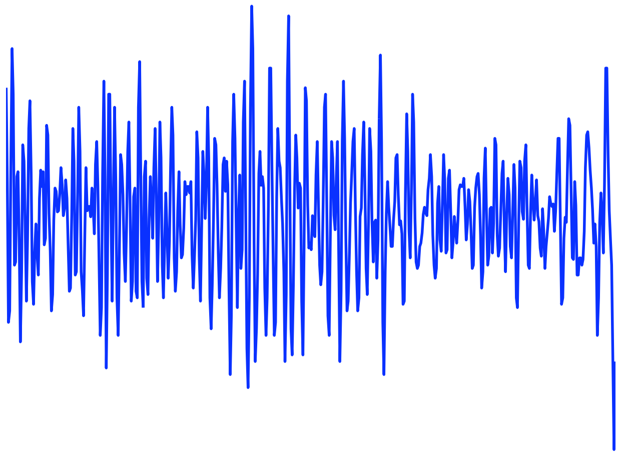
different

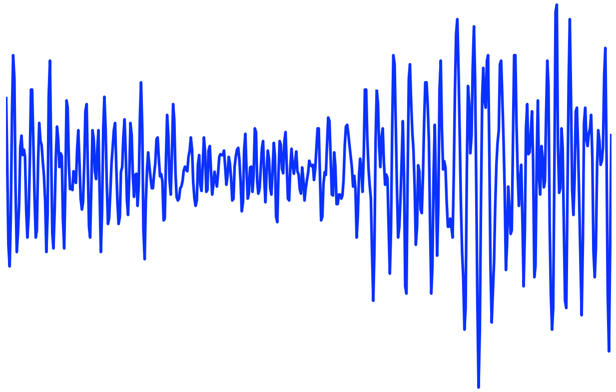
methods



How can we compare power spectra?

Quantitative analysis



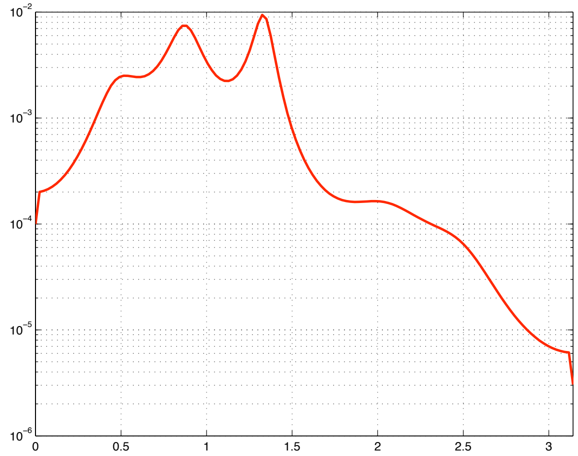


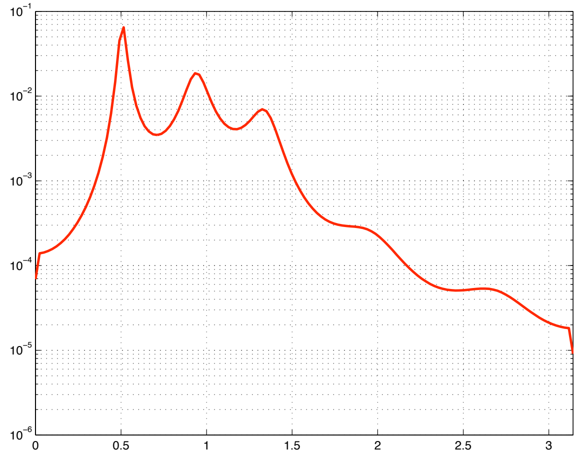
→

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?

Plan of the talk

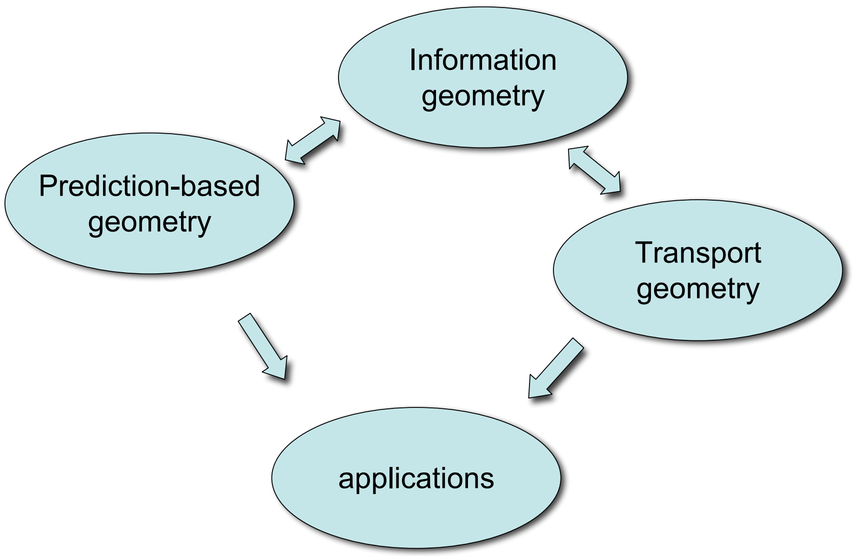
Metrics based on

prediction theory

some parallels with information geometry

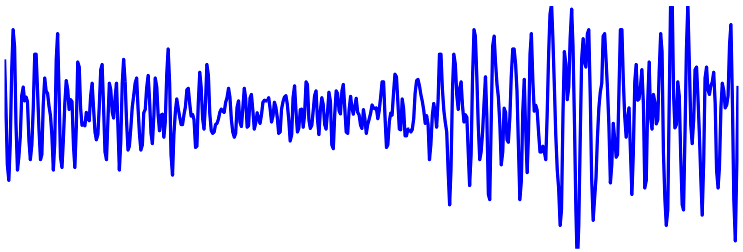
transport geometry

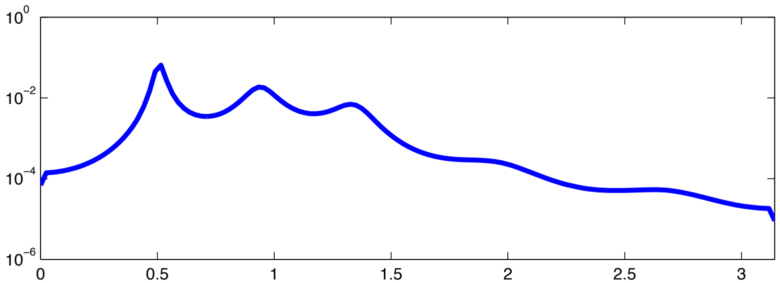
Case studies & applications



Setting

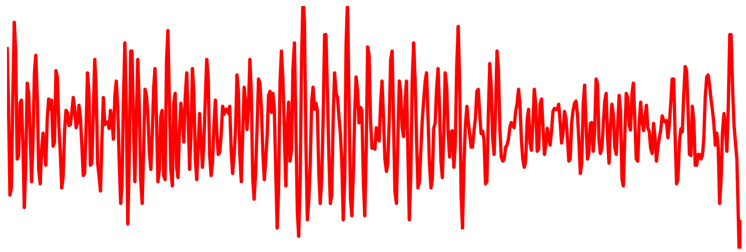
$\ldots u_{-1}, u_0, u_1, u_2, \ldots$

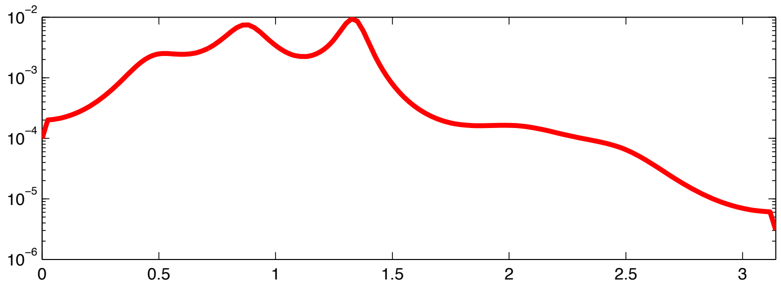




$f_1(\theta)$

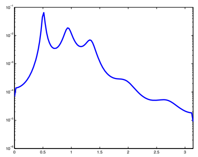
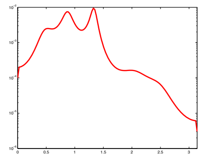
$\ldots u_{-1}, u_0, u_1, u_2, \ldots$





$f_2(\theta)$

What is it we would like to have?

distance ( , )

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

- metric
- meaningful & natural

candidates?

Kullback-Leibler, Bregman, Itakura-Saito, Makhoul,..

convex functionals

perceptual qualities

Linear prediction

One-step-ahead prediction: $u_{\text{present}} - \hat{u}_{\text{present}|\text{past}}$

with $\hat{u}_{\text{present}|\text{past}} := \sum_{\text{past}} \alpha_k u_k$

$$E\{|u_{\text{present}} - \hat{u}_{\text{present}|\text{past}}|^2\} = \text{variance of prediction error}$$

Szegö's theorem

One-step-ahead prediction:

$$\text{least error variance} = \exp\left\{\frac{1}{2\pi}\int \log f(\theta)d\theta\right\}$$

G. Szegö

it is a geometric mean...

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

Degradation of prediction error variance

Use f_2 to design a predictor (assuming $u_{f_2,\text{time}}$).

Then compare how this performs on $u_{f_1,\text{time}}$ against the optimal based on f_1.

$$\frac{\overbrace{E\{|u_{f_1,\text{present}} - \sum_{\text{past}} a_{f_2,\text{past}}\, u_{f_1,\text{past}}|^2\}}^{\text{degraded variance}} - \text{optimal variance}}{\text{optimal variance}} \geq 0$$

Degradation of prediction variance

$$\frac{\overbrace{E\{|u_{f_1,\text{present}} - \sum_{\text{past}} a_{f_2,\text{past}}\, u_{f_1,\text{past}}|^2\}}^{\text{degraded variance}}}{\text{optimal variance}} = \frac{\text{arithmetic mean of } \left(\frac{f_1}{f_2}\right)}{\text{geometric mean of } \left(\frac{f_1}{f_2}\right)}$$

$$= \frac{\left(\frac{1}{2\pi}\int (\frac{f_1}{f_2})d\theta\right)}{\exp\left(\frac{1}{2\pi}\int \log(\frac{f_1}{f_2})d\theta\right)}$$

arithmetic over *geometric* mean (≥ 1)

Riemannian metric

$$f_1 = f,$$
$$f_2 = f + \Delta$$

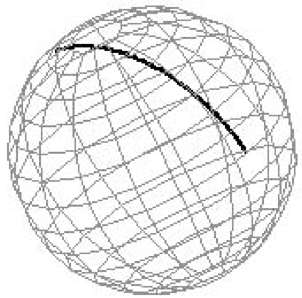
$$\frac{\overbrace{E\{|u_{f_1,\text{present}} - \sum_{\text{past}} a_{f_2,\text{past}}\, u_{f_1,\text{past}}|^2\}}^{\text{degraded variance}} - \text{optimal variance}}{\text{optimal variance}} \simeq$$

$$\delta(f, f+\Delta) = \frac{1}{2\pi}\int \left(\frac{\Delta}{f}\right)^2 d\theta - \left(\frac{1}{2\pi}\int \left(\frac{\Delta}{f}\right) d\theta\right)^2$$

variance-like: (mean square) - (arithmetic-mean)^2

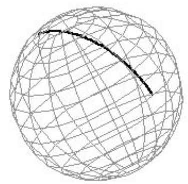
Geodesics

Paths $f_{\mathbf{r}}$ ($\mathbf{r} \in [0, 1]$) between f_0, f_1 of minimal length $\int_0^1 \sqrt{\delta\left(f_{\mathbf{r}}, f_{\mathbf{r}+d\mathbf{r}}\right)}$



each point represents a different power spectral density

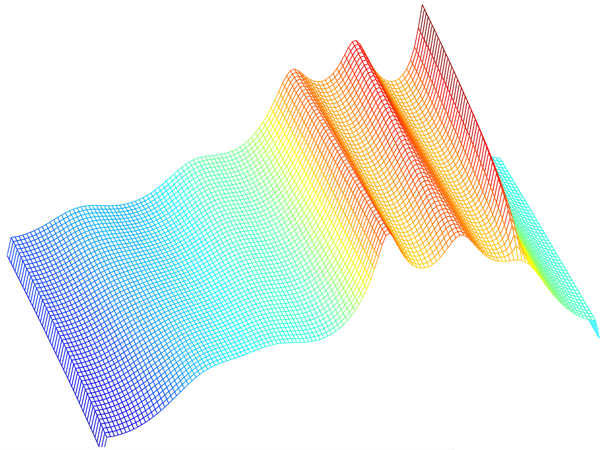
Geodesics

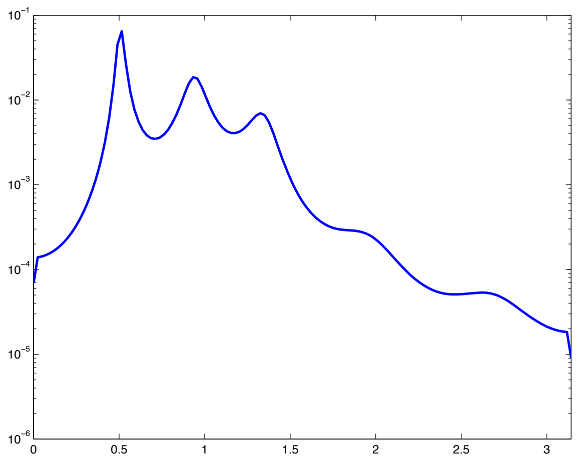


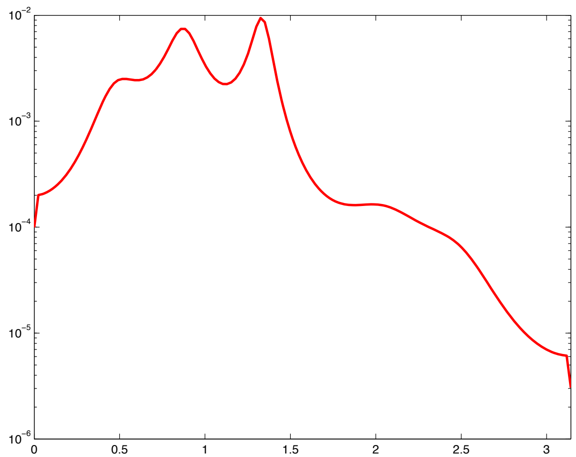
The geodesics are exponential families:

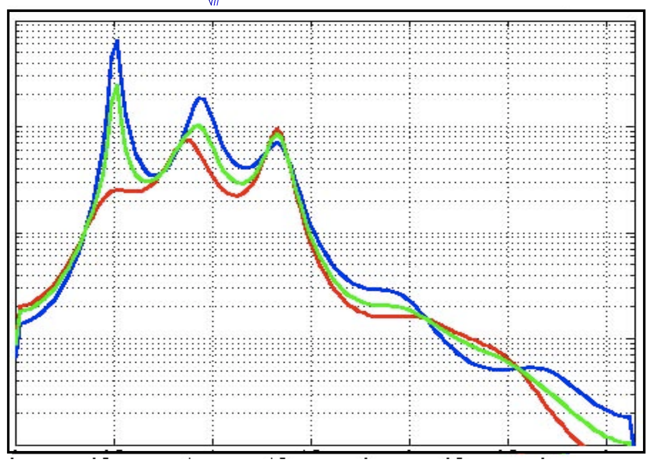
$$f_{\mathbf{r}} = f_0 \left(\frac{f_1}{f_0} \right)^{\mathbf{r}}, \quad \mathbf{r} \in [0, 1]$$

$$= \exp\{ \, (1 - \mathbf{r}) \log(f_0) + \mathbf{r} \log(f_1) \, \}$$



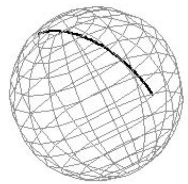






morphing

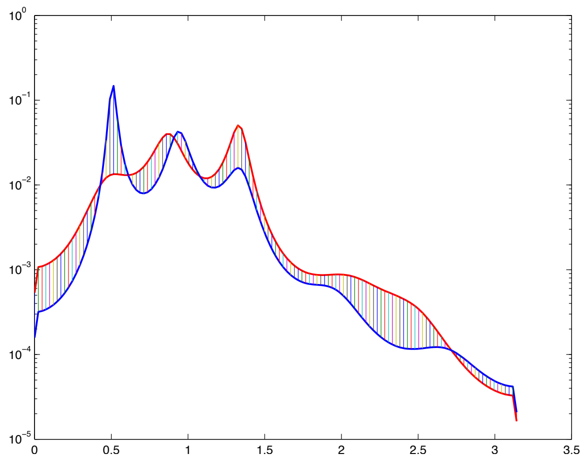
Geodesic distance: metric

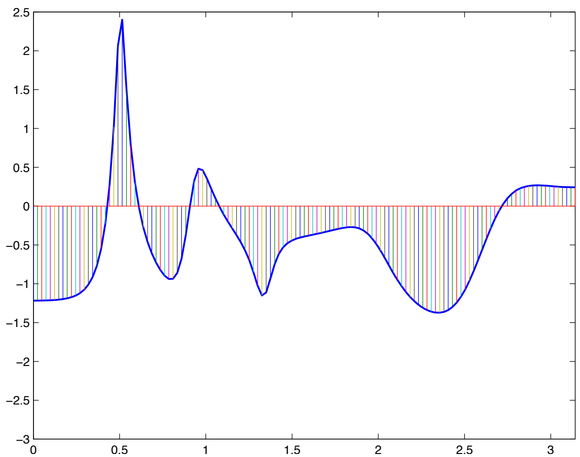


The path-length is

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

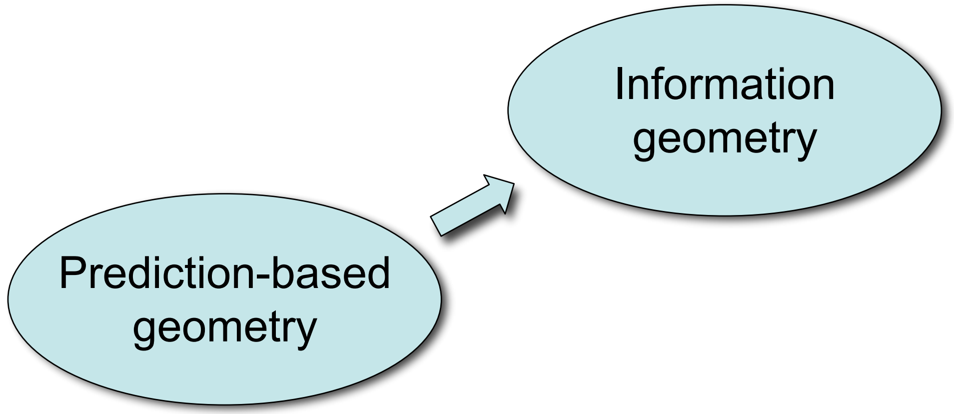
variance-like distance on logarithms: (mean square) - (arithmetic-mean)^2
scale-insensitive, “shape” recognizer





In IEEE Trans. on Signal Processing, Aug. 2007

$$\log\frac{f_1}{f_0} \quad = \quad \log(f_1) - \log(f_0)$$


Information
geometry
Prediction-based
geometry

Information geometry – *parallels*

$f \rightsquigarrow \mathbf{p}$: probability density

$$I = E_{\mathbf{p}}\{(\partial_\lambda \log \mathbf{p}_\lambda)^2\} \quad \delta\lambda^2$$

R. Fisher

Fisher information metric

$$I = \sum \frac{\Delta^2}{\mathbf{p}}$$

C.R. Rao

Information geometry – *parallels*

R. Leibler

Expected "message-length increase":

$$H(\mathbf{p}_1|\mathbf{p}_0) = \left(-\sum \mathbf{p}_1 \log(\mathbf{p}_0)\right) - \left(-\sum \mathbf{p}_1 \log(\mathbf{p}_1)\right)$$

S. Kullback

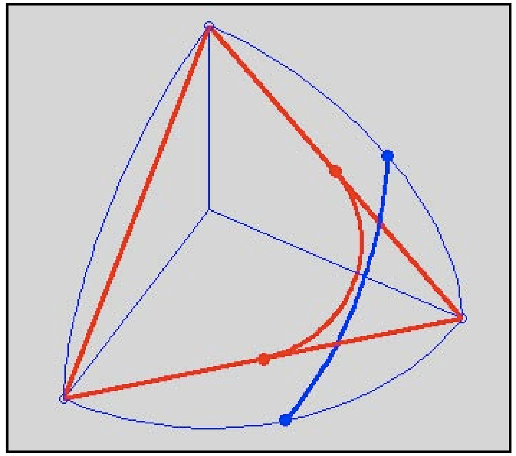
Fisher information metric

$\mathbf{p}_0 = \mathbf{p}$

$\mathbf{p}_1 = \mathbf{p} + \Delta$

$$I = \sum \frac{\Delta^2}{\mathbf{p}}$$

Information geometry – *parallels*

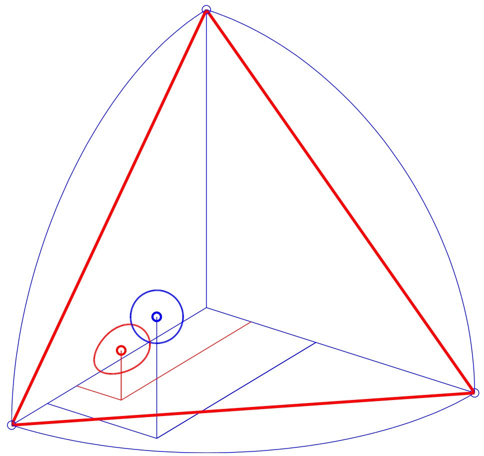


Geodesics: great circles

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

$$\begin{pmatrix} p(1) \\ p(2) \\ p(3) \end{pmatrix} \mapsto \begin{pmatrix} \sqrt{p(1)} \\ \sqrt{p(2)} \\ \sqrt{p(3)} \end{pmatrix}$$

Geodesic distance: Arclength
Battacharyya distance



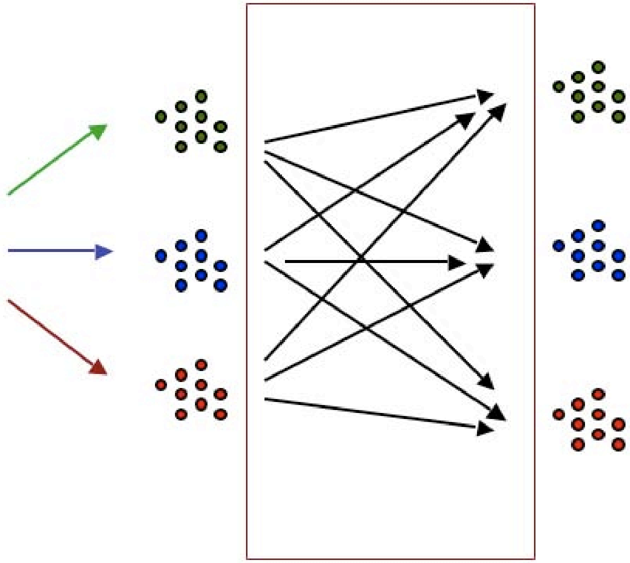
Information vs. prediction-based

$$\sum \frac{\Delta^2}{\mathbf{p}} \quad \text{vs.} \quad \int \left(\frac{\Delta}{\mathbf{f}}\right)^2 - \left(\int \frac{\Delta}{\mathbf{f}}\right)^2$$

$$\mathbf{p} \mapsto \sqrt{\mathbf{p}} \quad \text{vs.} \quad \mathbf{f} \mapsto \log \mathbf{f}$$

great circles vs. **logarithmic families**

Information geometry – *parallels*



Ability to differentiate decreases

$$\begin{pmatrix} p(1) \\ p(2) \\ p(3) \end{pmatrix} \mapsto \mathrm{M} \begin{pmatrix} p(1) \\ p(2) \\ p(3) \end{pmatrix}$$

Chentsov's theorem:

Stochastic maps are contractive

under *Fisher metric*

and

Fisher metric is the unique Riemannian metric with this property

What is the analog for power spectra?

addititive noise

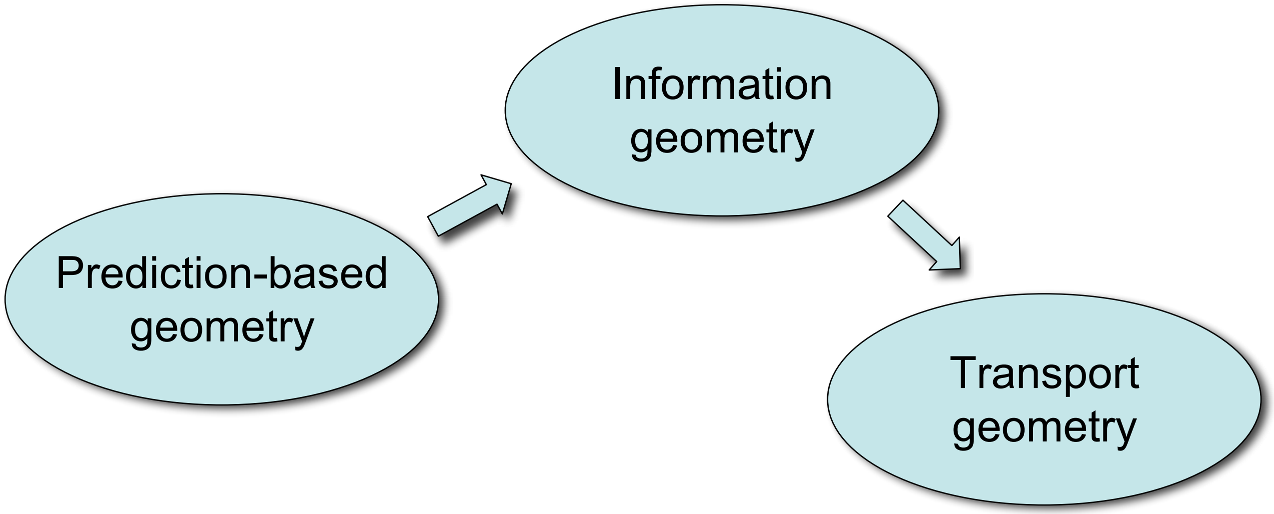
$$f \;\; \mapsto f + f_{\text{noise}}$$

multiplicative noise

$$f \;\; \mapsto f \star f_{\text{noise}}$$

continuity of moments (second-order statistics)

$$f \;\; \mapsto \text{integrals of } f$$

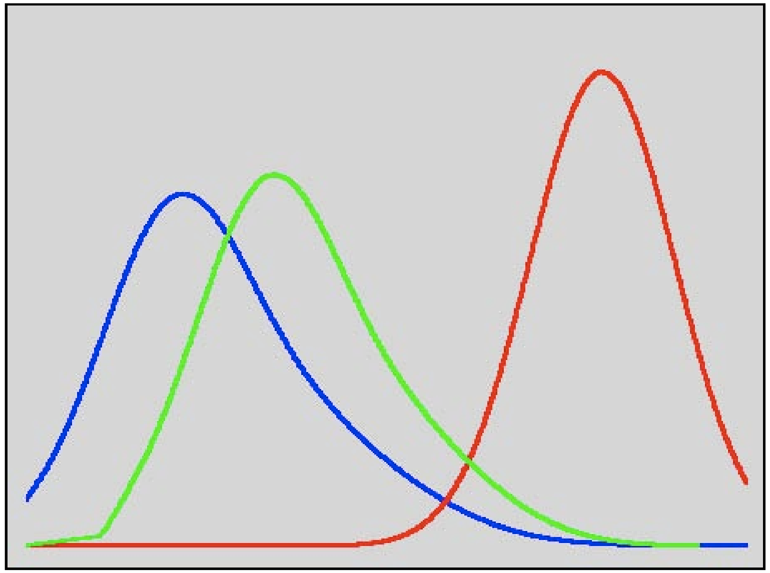

Information geometry
Prediction-based geometry
Transport geometry

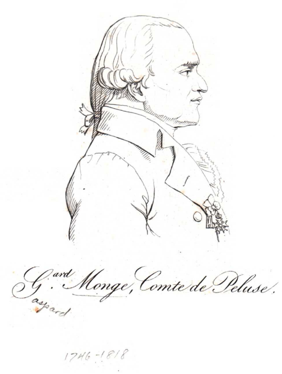
Transport geometry

Monge-Kantorovich problem

minimize cost of transferring mass

$$\int \text{cost}(x \to y) \;\times\; \text{mass}(dx, dy)$$





G. Monge

L. Kantorovich

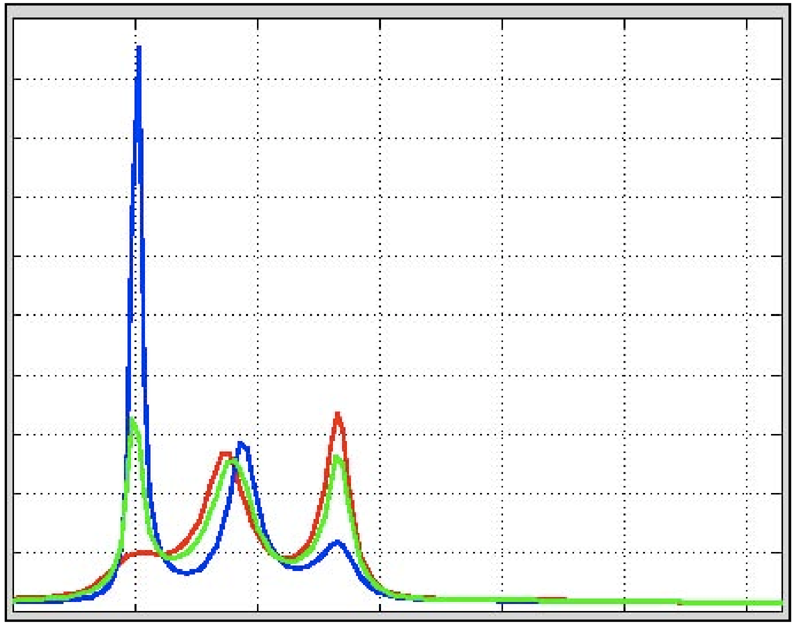
Transport for power spectra

Transport-based metric

distances do not increase

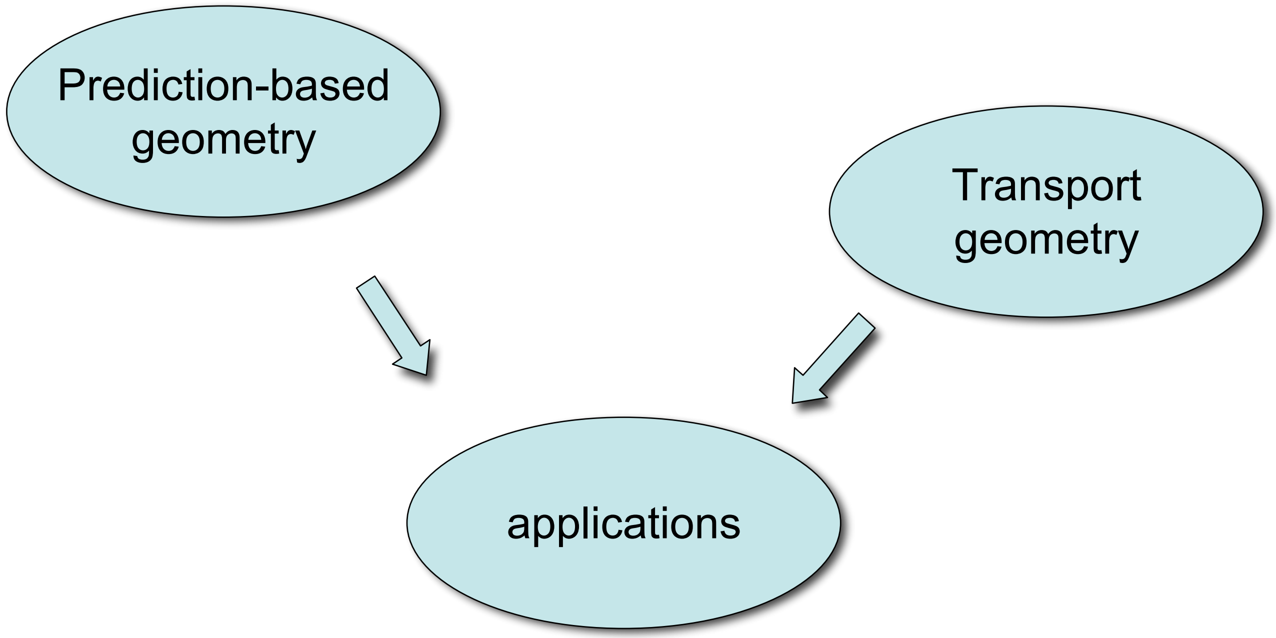
under additive noise
and multiplicative noise
with power ≤ 1

$+$ continuity of statistics

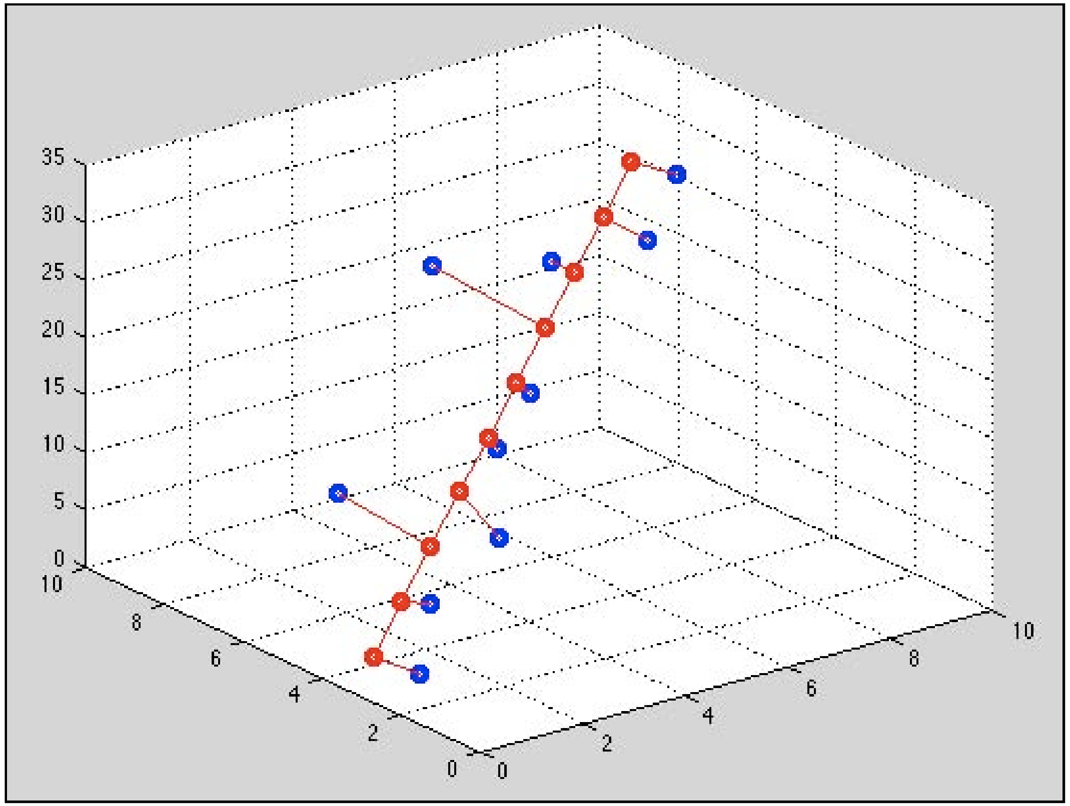


$$\mathbf{metric} \;=\; \min\left(\text{cost of transport}(\hat{f}_0, \hat{f}_1) \;+\; \text{normalization}\right)$$

with Johan Karlsson (KTH) & Mir Shahrouz Takyar


Prediction-based geometry
Transport geometry
applications

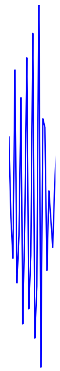
Fitting geodesics

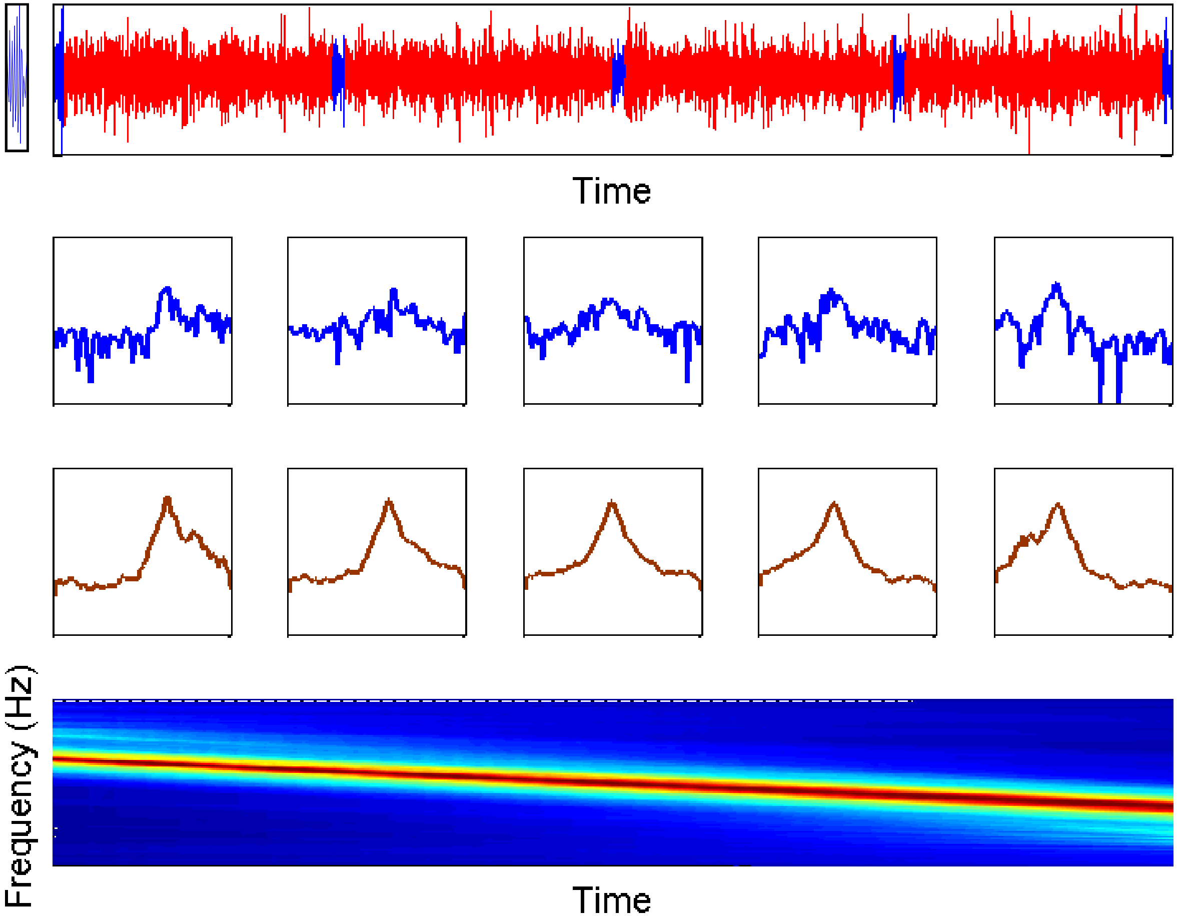


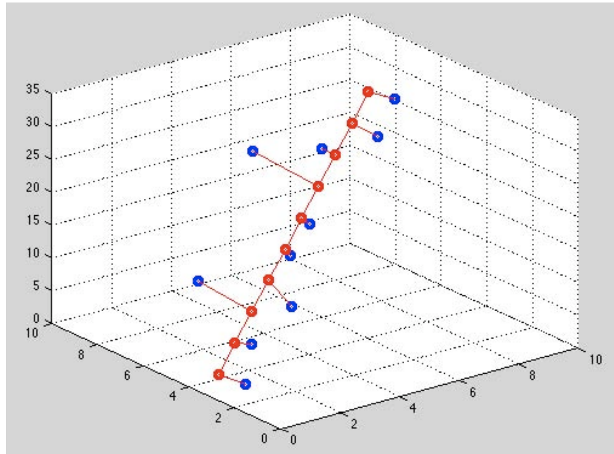
K.F. Gauss

Least squares: The theory of motion of heavenly bodies, Gauss, K.F.

Tracking with geodesics





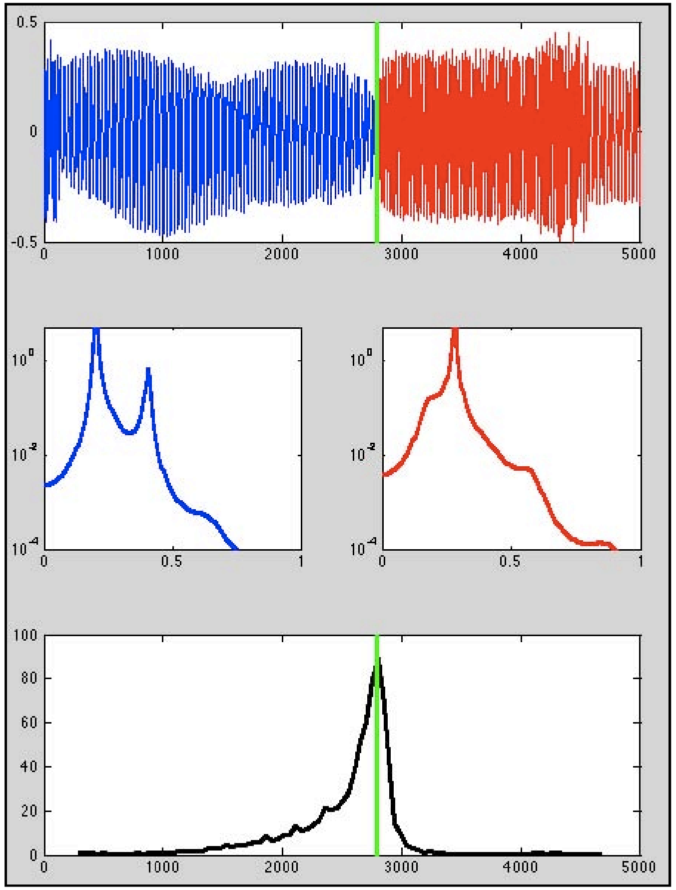


with Xianhua Jiang

Voice & sounds

John Weissmuller's MGM Tarzan Yell

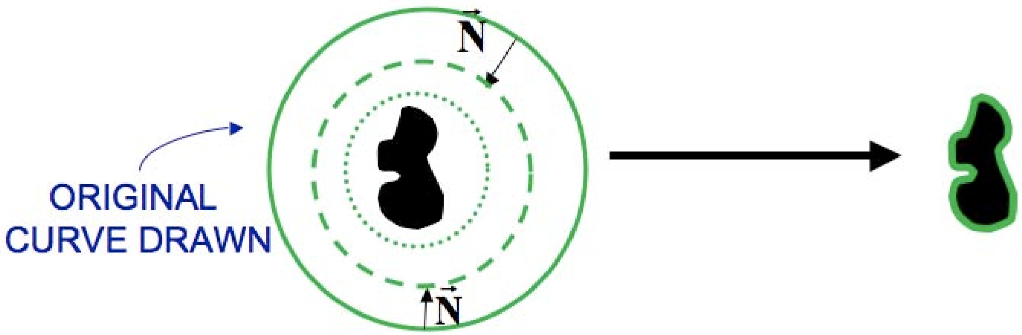
http://www.complxmind.com

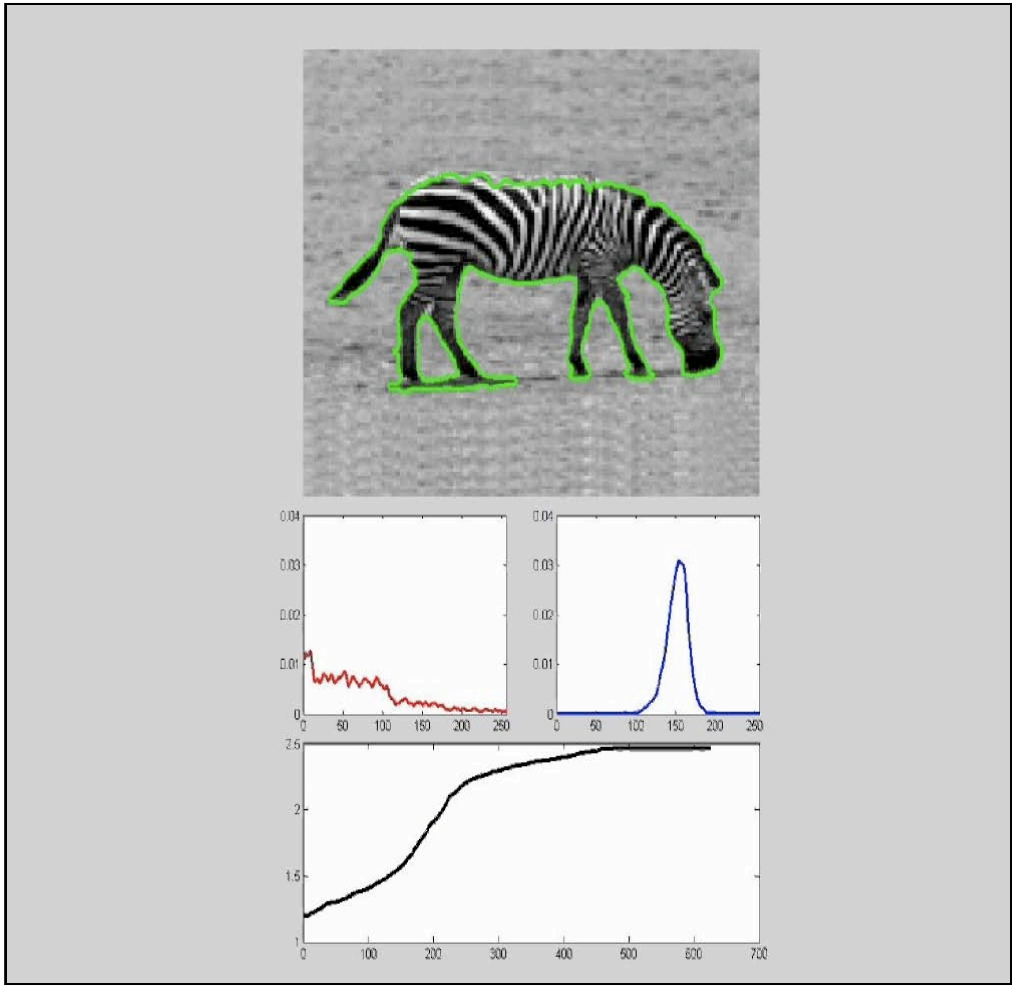


Images & more

Geometric active contours

$$\frac{\partial}{\partial t}\text{Curve} = \nabla_{\text{Curve}} \ \text{metric}(f_{\text{inside}}, f_{\text{outside}})$$





with Romeil Sandhu and Allen Tannenbaum

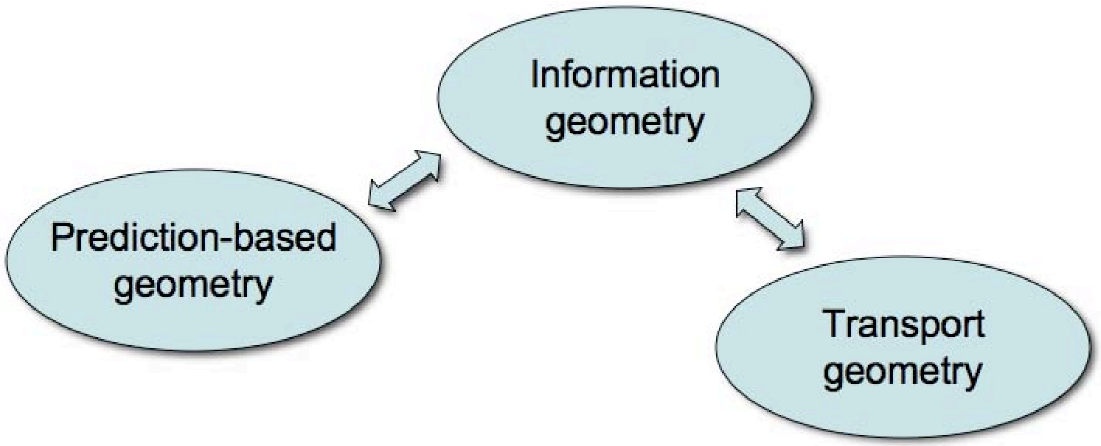
Images & more

with Romeil Sandhu and Allen Tannenbaum

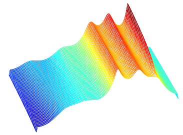
Concluding thoughts

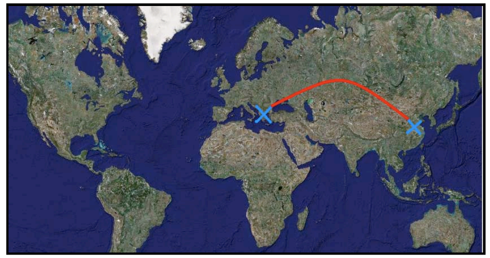
Metrics
in spectral analysis

- Operational significance
- Effect of natural transformations



Thank you for your attention





thanks to

Xianhua Jiang | Johan Karlsson | Romeil Sandhu | Mir Shahrouz Takyar




Allen Tannenbaum & Anders Lindquist

National Science Foundation, AFOSR, and Hermes-Luh endowment