Variational approach to time series analysis

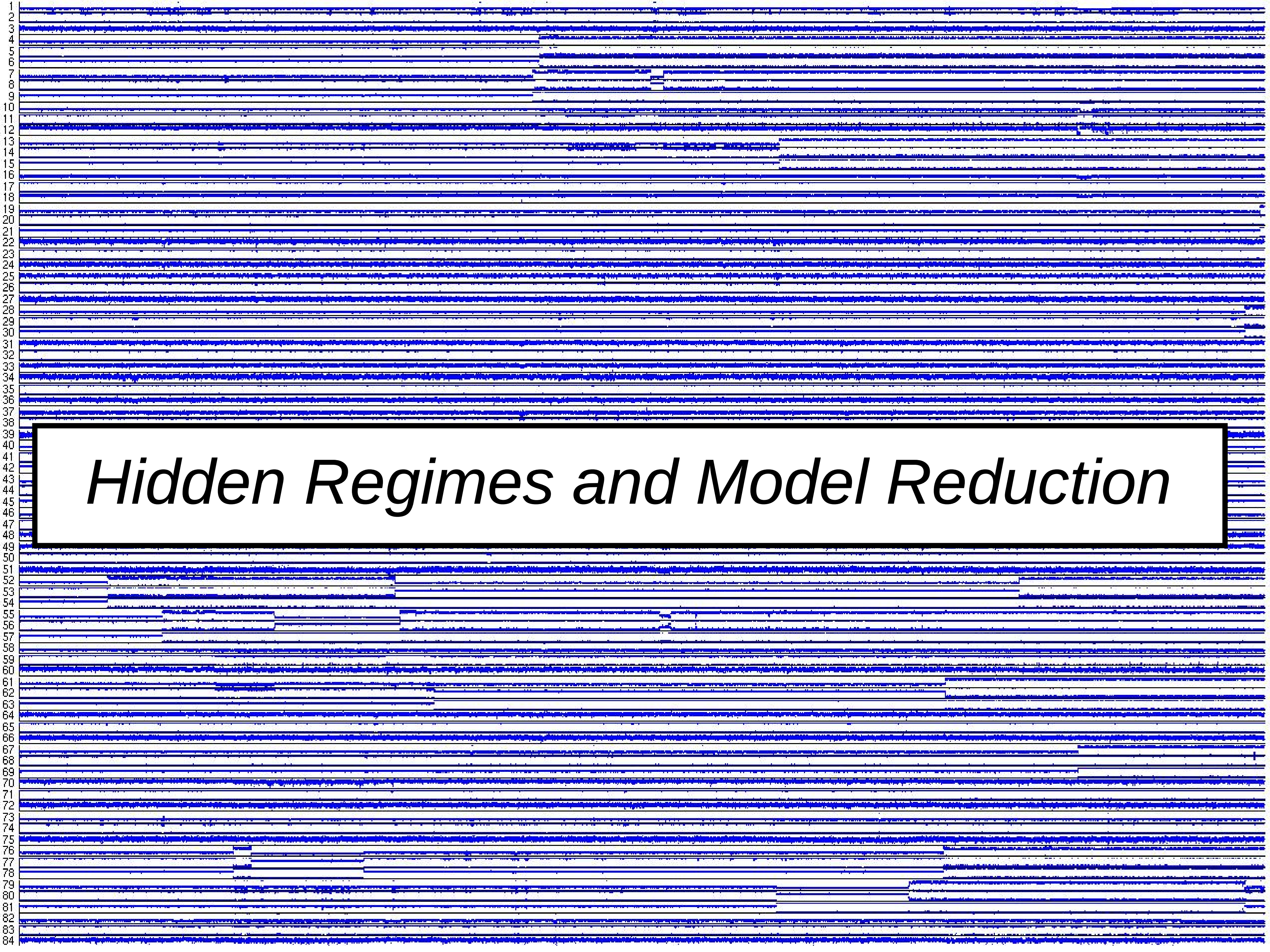
Illia Horenko

Scientific Computing
Institute of Mathematics
Freie Universität Berlin (FU)

DFG Research Center MATHEON
„Mathematics in key technologies“

- <u>*Hidden Regimes and Model Reduction*</u>
  - overwiev of the standard clustering methods, their problems
  - *variational approach*: regularized model distance functional
  - *finite element* approach to *clustering problem*
  - *manifold clustering*
  - *toy examples*
- <u>*Applications*:</u>
  - analysis of weather data (*regimes*, *prediction*)
  - computational finance (*portfolio theory*)
  - compression of 3D *turbulence simulation data*

Hidden Regimes and Model Reduction



Model Distance Functional

Let $x(t) : \mathbf{R}^1 \longrightarrow \Psi \subset \mathbf{R}^n$ be the *observed process* $t \in [0, T]$

Define $\mathbf{K}$ local models by a *model distance functional*:

$$g\left(x, \theta_i\right) \quad : \quad \Psi \times \Omega \longrightarrow [0, \bar{g}] \,, \quad 0 < \bar{g} < +\infty,$$

$$\theta_1, \ldots, \theta_{\mathbf{K}} \in \Omega \subset \mathbf{R}^d$$

Examples

- *Geometrical clustering*: $\theta_i \in \Psi$ - cluster centers

$$g\left(x, \theta_i\right) \quad = \quad \| \, x - \theta_i \, \|^2,$$

Model Distance Functional

Let $x(t) : \mathbf{R}^1 \to \Psi \subset \mathbf{R}^n$ be the *observed process* $t \in [0, T]$

Define $\mathbf{K}$ local models by a *model distance functional*:

$$g(x, \theta_i) \quad : \quad \Psi \times \Omega \to [0, \bar{g}], \quad 0 < \bar{g} < +\infty,$$

$$\theta_1, \ldots, \theta_{\mathbf{K}} \in \Omega \subset \mathbf{R}^d$$

Examples

- *Geometrical clustering*: $\theta_i \in \Psi$ - cluster centers

$$g(x, \theta_i) = \| x - \theta_i \|^2,$$

- *Gaussian clustering*: $\theta_i = (\mu_i, \Sigma_i)$ - Gaussian parameters

$$g(x, \theta_i) = \| x - \mu_i \|^2_{\Sigma_i^{-1}}$$

Clustering Problem: definition

What is "clustering"?

Find $\Gamma(t) = (\gamma_1(t), \ldots, \gamma_{\mathbf{K}}(t))$ such that for each t :

$$\sum_{i=1}^{\mathbf{K}} \gamma_i(t) g(x, \theta_i) \quad \longrightarrow \quad \min_{\Gamma(t), \Theta},$$

subjected to constraints:

$$\begin{aligned} \sum_{i=1}^{\mathbf{K}} \gamma_i(t) &= 1, \quad \forall t \in [0, T] \\ \gamma_i(t) &\geq 0, \quad \forall t \in [0, T], i = 1, \ldots, \mathbf{K} \end{aligned}$$

Averaged Clustering Functional

Find $\Gamma(t) = (\gamma_1(t), \ldots, \gamma_{\mathbf{K}}(t))$ such that for each t :

$$\mathbf{L}(\Theta, \Gamma(t)) = \int_0^T \sum_{i=1}^{\mathbf{K}} \gamma_i(t) g\left(x_t, \theta_i\right) \rightarrow \min_{\Gamma(t), \Theta},$$

subjected to constraints:

$$\sum_{i=1}^{\mathbf{K}} \gamma_i(t) = 1, \quad \forall t \in [0, T]$$

$$\gamma_i(t) \geq 0, \quad \forall t \in [0, T], i = 1, \ldots, \mathbf{K}$$

Numerical Method: Subspace Iteration (splitting scheme)

No global convergence (non-convex optimization, simulated annealing)

K-Means clustering: problems

Geometrical distance: $\theta_i \in \Psi$ - *time-independent* cluster centers

$$g(x, \theta_i) = \| x - \theta_i \|^2,$$

$$t_j, j = 1, \ldots, n \in [0, T]$$

$$\sum_{i=1}^{\mathbf{K}} \sum_{j=1}^{n} \gamma_i(t_j) \| x_{t_j} - \theta_i \|^2 \rightarrow \min_{\Gamma(t), \Theta}$$

(Bezdek1981, Höppner et.al. 1999)

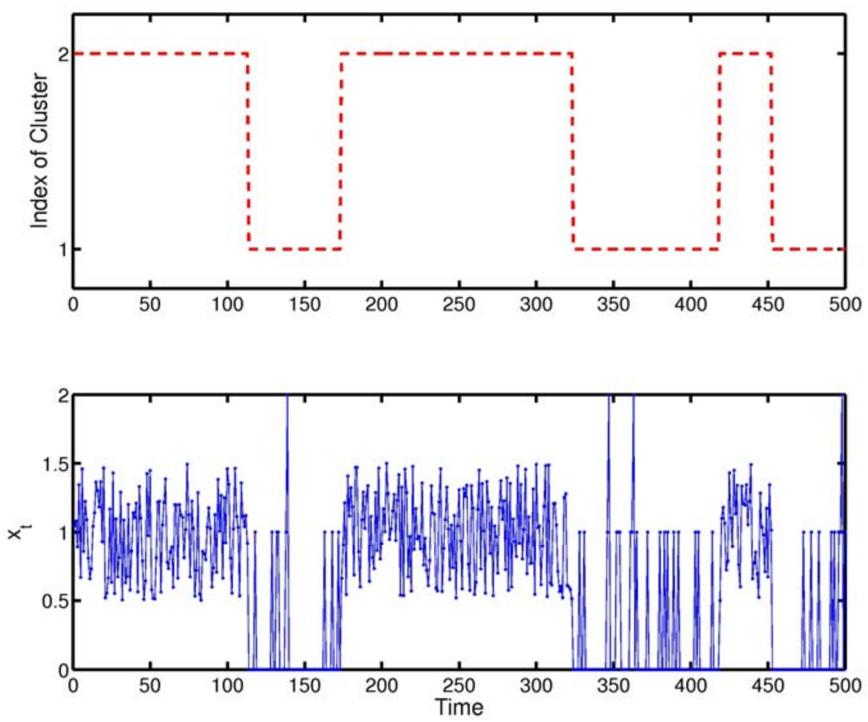
Iteration number (l):

$$\gamma_i^{(l)}(t_j) = \begin{cases} 1 & i = \arg\min \| x_{t_j} - \theta_i^{(l-1)} \|^2; \\ 0 & \text{otherwise}, \end{cases}$$

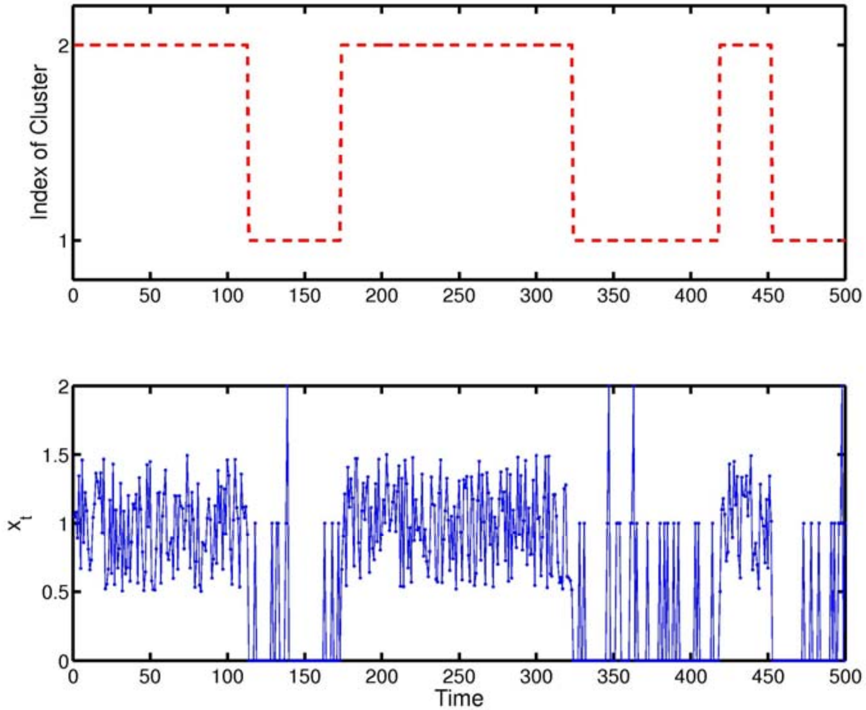
$$\theta_i^{(l)} = \frac{\sum_{j=1}^{n} \gamma_i^{(l)}(t_j) x_{t_j}}{\sum_{j=1}^{n} \gamma_i^{(l)}(t_j)}.$$

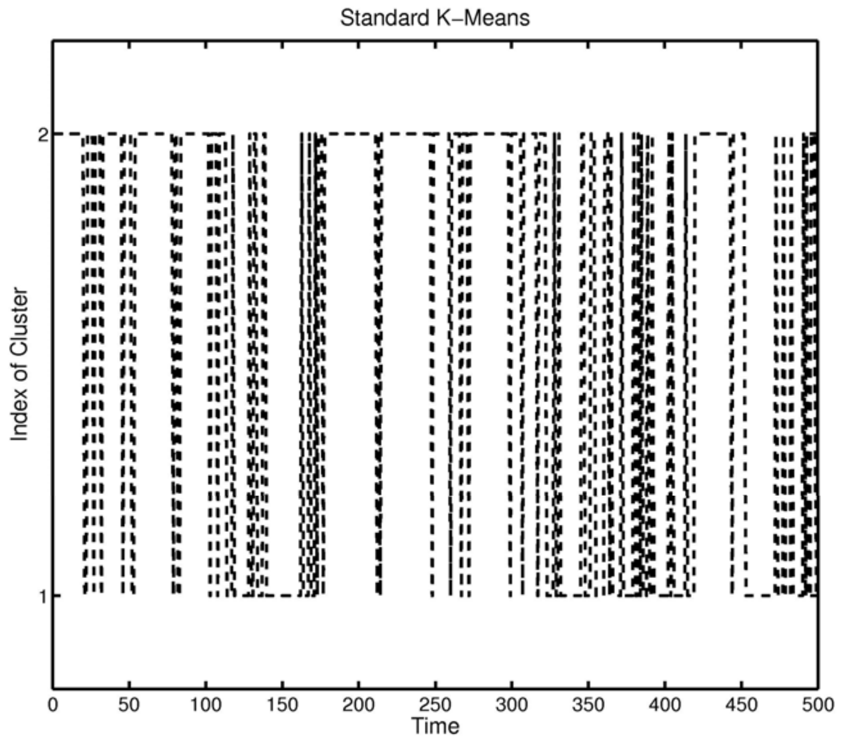
*"Sharp" assigment : **problem for overlapping data***

K-Means clustering: Toy Example I



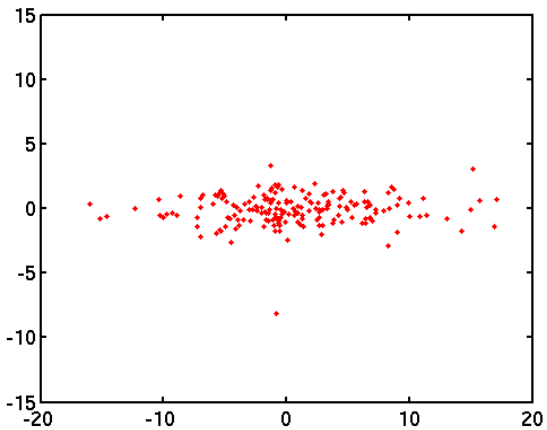
K-Means clustering: Toy Example I



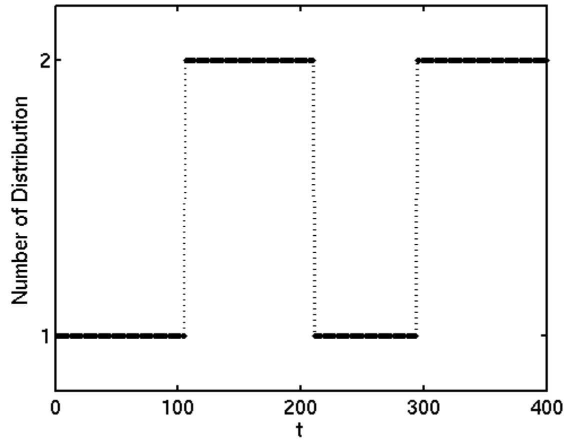


K-Means clustering: Toy Example II

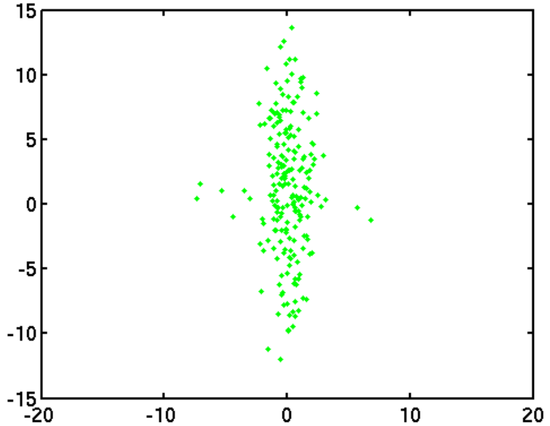
Distribution 1



Switching between the distributions



Distribution 2



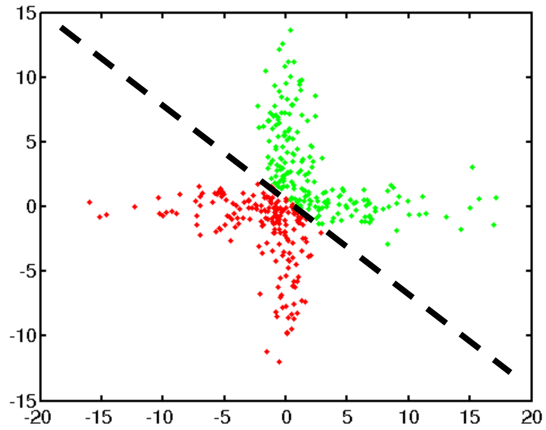
$$\mathbf{L}(\Theta,\Gamma(t)) = \int_0^T \sum_{i=1}^{\mathbf{K}} \gamma_i(t) g\left(x_t,\theta_i\right) \rightarrow \min_{\Gamma(t),\Theta},$$

$$g\left(x,\theta_i\right) = \| x-\theta_i \|^2,$$

K-Means-Algorithm

K-Means clustering: Toy Example II

colouring from
geometrical clustering



$$\mathbf{L}(\Theta, \Gamma(t)) = \int_0^T \sum_{i=1}^{\mathbf{K}} \gamma_i(t) g(x_t, \theta_i) \rightarrow \min_{\Gamma(t), \Theta},$$

$$g(x, \theta_i) = \| x - \theta_i \|^2,$$

K-Means-Algorithm

Problems:

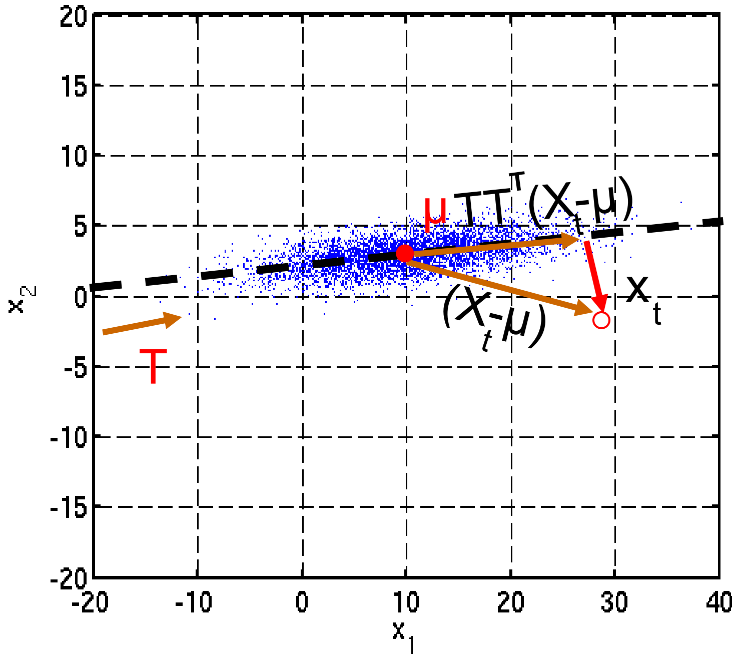
1. *Euclidean distance may be not appropriate*
2. *geometrical clustering gets no use of temporal information*

Manifold Clustering (H.06-07)

Problem 1: $\mathbf{L}\left(\gamma_i(t), \mathbf{T}_i, \mu_i\right) \rightarrow \min$

$$g\left(x, \theta_i\right) \quad = \quad \| \; x - \mathcal{T}_i \mathcal{T}_i^{\mathbf{T}} x \; \|^2$$

$\theta_i = \mathcal{T}_i \in \mathbf{R}^{nxm}$ ***m << n***

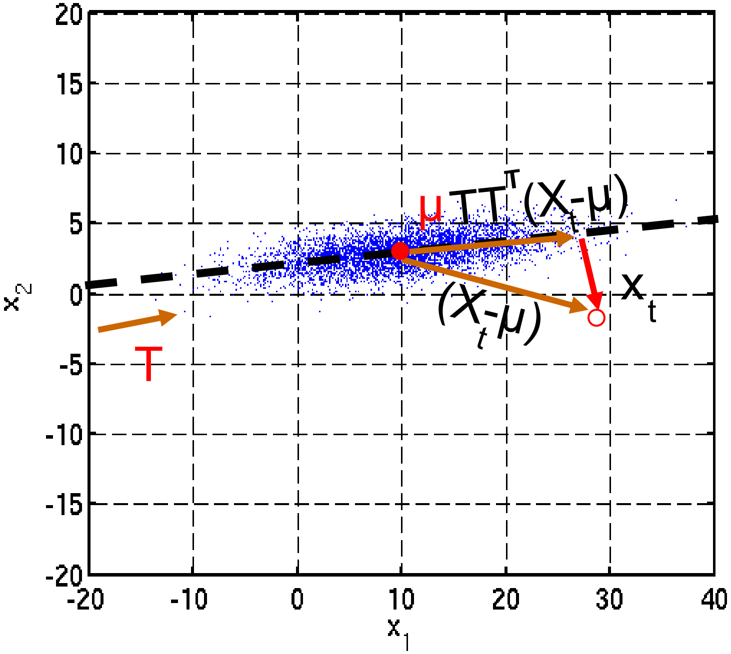


Manifold Clustering (H.06-07)

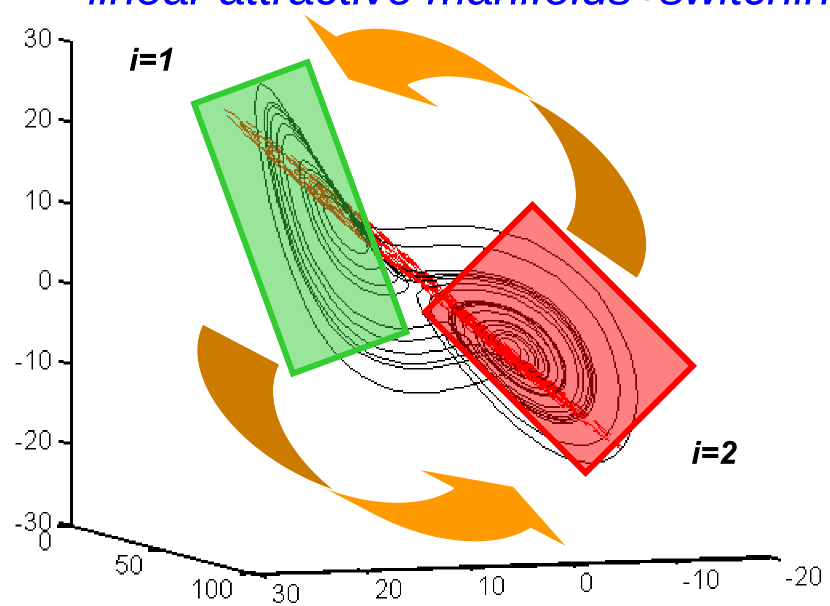
Problem 1: $\mathbf{L}\left(\gamma_i(t), \mathbf{T}_i, \mu_i\right) \longrightarrow \min$

$$g\left(x, \theta_i\right) \;=\; \| \, x - \mathcal{T}_i \mathcal{T}_i^{\mathbf{T}} x \, \|^2$$

$\theta_i = \mathcal{T}_i \in \mathbf{R}^{nxm}$ ***m << n***



***Idea**: Essential manifold*
can be appproximated by
linear attractive manifolds+switching



H./Schmidt-Ehrenberg/Schütte 06
H. 07

Fuzzy C-Means Algorithm

Geometrical distance: $\theta_i \in \Psi$ - *time-independent* cluster centers,

introduce the *"fuzzifier"* $m > 1$

$$\sum_{i=1}^{\mathbf{K}} \sum_{j=1}^{n} \gamma_i^m(t_j) \parallel x_{t_j} - \theta_i \parallel^2 \quad \longrightarrow \quad \min_{\Gamma(t),\Theta} \qquad \textit{(Bezdek1987)}$$

$$\mathbf{I}_{x_{t_j}} = \{p \in \{1,\ldots,\mathbf{K}\} | \parallel x_{t_j} - \theta_p^{(l-1)} \parallel^2 = 0\}$$

$$\gamma_i^{(l)}(t_j) = \begin{cases} \dfrac{1}{\sum_{p=1}^{\mathbf{K}} \left(\frac{\| x_{t_j} - \theta_i^{(l-1)} \|^2}{\| x_{t_j} - \theta_p^{(l-1)} \|^2} \right)^{\frac{1}{m-1}}} & \text{if } \mathbf{I}_{x_{t_j}} \text{ is empty,} \\ \sum_{r \in \mathbf{I}_{x_{t_j}}} \gamma_r^{(l)}(t_j) = 1 & \text{if } \mathbf{I}_{x_{t_j}} \text{ is not empty, } i \in \mathbf{I}_{x_{t_j}}, \\ 0 & \text{if } \mathbf{I}_{x_{t_j}} \text{ is not empty, } i \notin \mathbf{I}_{x_{t_j}}, \end{cases}$$

$$\theta_i^{(l)} = \frac{\sum_{j=1}^{n} \gamma_i^{(l)}(t_j) x_{t_j}}{\sum_{j=1}^{n} \gamma_i^{(l)}(t_j)}.$$

Fuzzy C-Means Algorithm

Exercise 1: *proof that this formulas are true*

$$\gamma_i^{(l)}(t_j) = \begin{cases} \dfrac{1}{\sum_{p=1}^{K}\left(\frac{\|x_{t_j}-\theta_i^{(l-1)}\|^2}{\|x_{t_j}-\theta_p^{(l-1)}\|^2}\right)^{\frac{1}{m-1}}} & \text{if } \mathbf{I}_{x_{t_j}} \text{ is empty,} \\ \sum_{r\in \mathbf{I}_{x_{t_j}}} \gamma_r^{(l)}(t_j) = 1 & \text{if } \mathbf{I}_{x_{t_j}} \text{ is not empty, } i \in \mathbf{I}_{x_{t_j}}, \\ 0 & \text{if } \mathbf{I}_{x_{t_j}} \text{ is not empty, } i \notin \mathbf{I}_{x_{t_j}}, \end{cases}$$

$$\theta_i^{(l)} = \frac{\sum_{j=1}^{n} \gamma_i^{(l)}(t_j) x_{t_j}}{\sum_{j=1}^{n} \gamma_i^{(l)}(t_j)}.$$

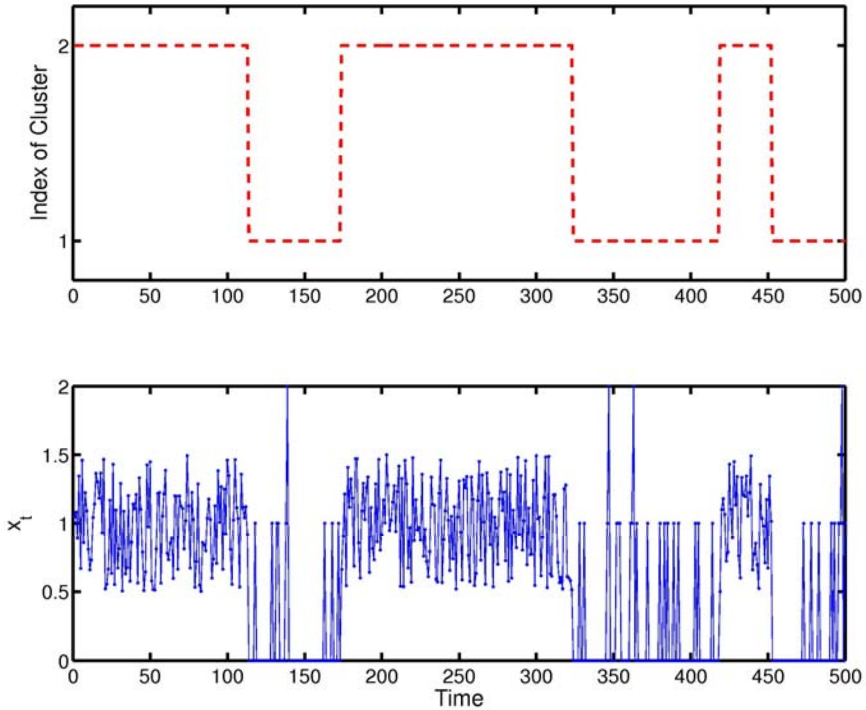
Fuzzy C-Means Algorithm

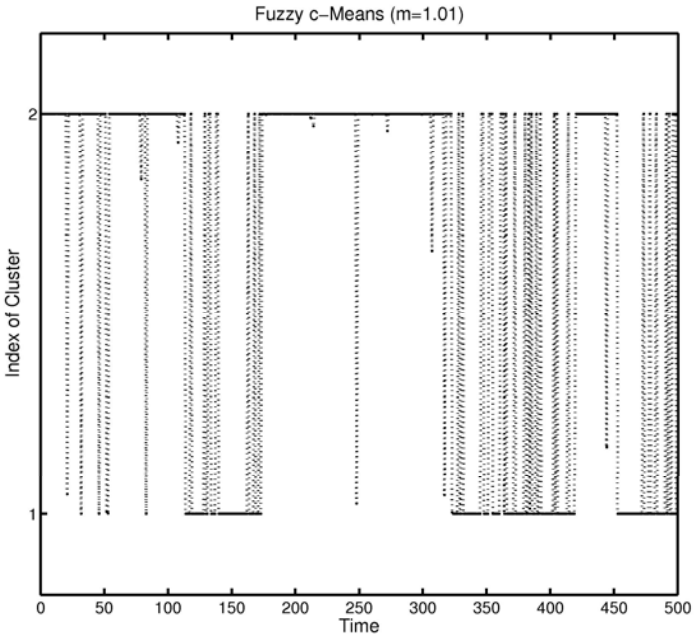
Geometrical distance: $\theta_i \in \Psi$ - *time-independent* cluster centers,

introduce the *"fuzzifier"* $m > 1$

$$\sum_{i=1}^{K} \sum_{j=1}^{n} \gamma_i^m(t_j) \parallel x_{t_j} - \theta_i \parallel^2 \longrightarrow \min_{\Gamma(t),\Theta}$$

(Bezdek1987)





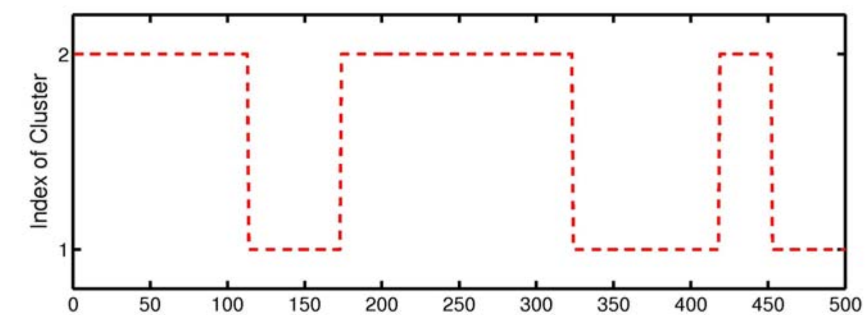
Fuzzy C-Means Algorithm

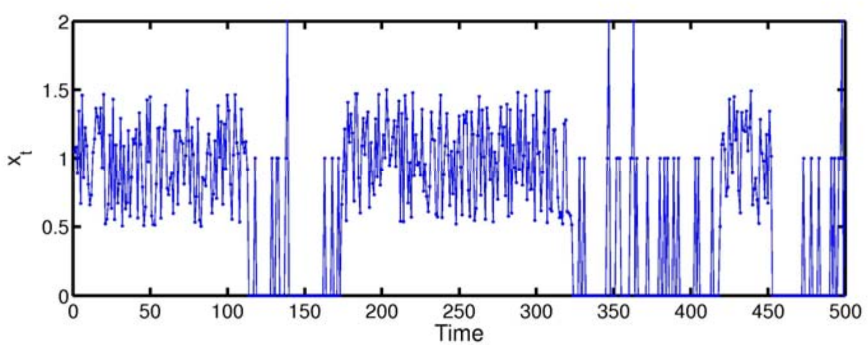
Geometrical distance: $\theta_i \in \Psi$ - *time-independent* cluster centers,

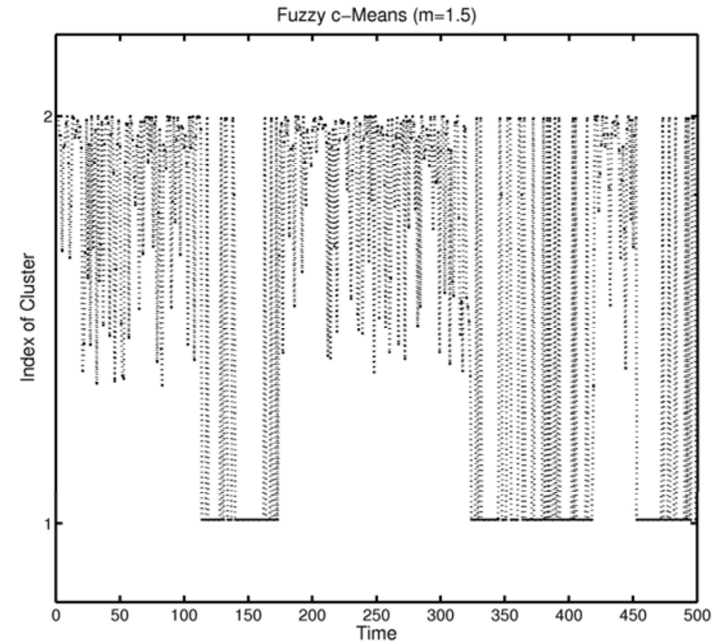
introduce the *"fuzzifier"* $m > 1$

$$\sum_{i=1}^{K}\sum_{j=1}^{n}\gamma_i^m(t_j)\,\|\,x_{t_j}-\theta_i\,\|^2 \;\longrightarrow\; \min_{\Gamma(t),\Theta}$$

(Bezdek1987)







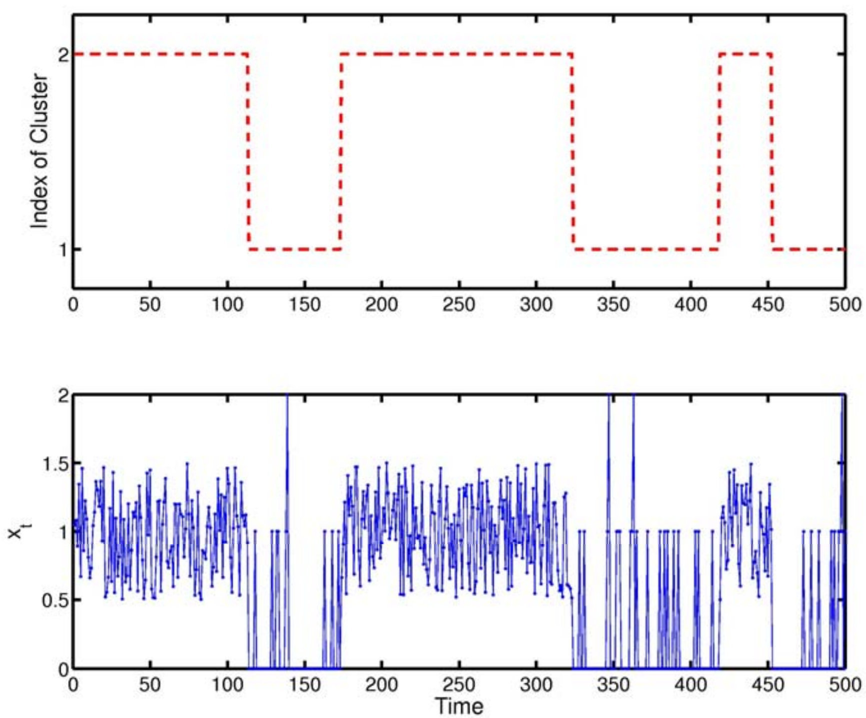
Fuzzy C-Means Algorithm

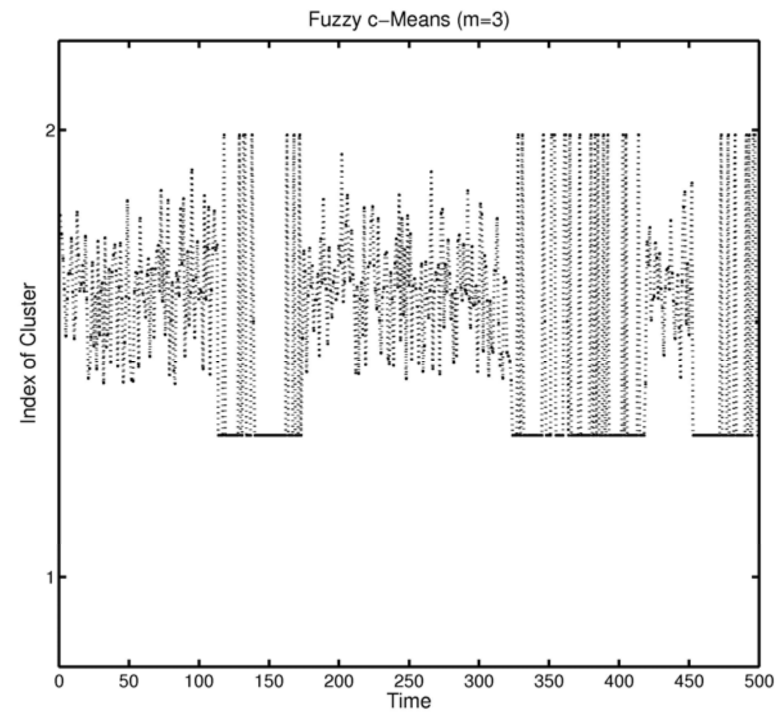
Geometrical distance: $\theta_i \in \Psi$ - *time-independent* cluster centers,

introduce the *"fuzzifier"* $m > 1$

$$\sum_{i=1}^{K}\sum_{j=1}^{n}\gamma_i^m(t_j)\,\|\,x_{t_j}-\theta_i\,\|^2 \;\longrightarrow\; \min_{\Gamma(t),\Theta}$$

(Bezdek1987)





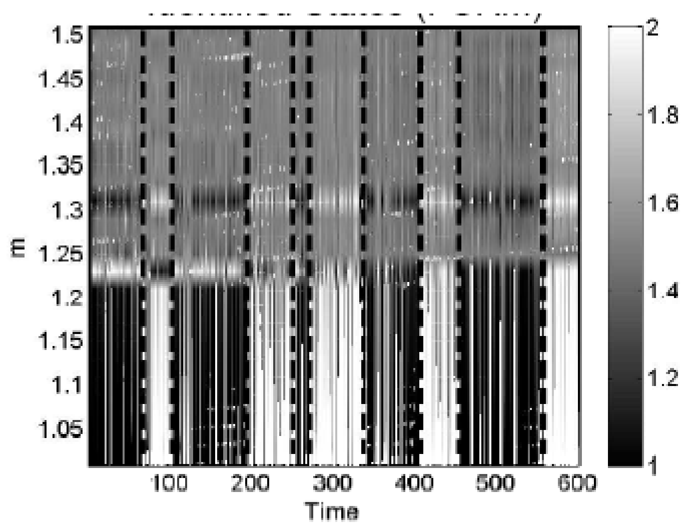
Fuzzy C-Means Algorithm

Geometrical distance: $\theta_i \in \Psi$ - *time-independent* cluster centers,

introduce the *"fuzzifier"* $m > 1$

$$\sum_{i=1}^{K} \sum_{j=1}^{n} \gamma_i^m(t_j) \parallel x_{t_j} - \theta_i \parallel^2 \quad \longrightarrow \quad \min_{\Gamma(t),\Theta}$$

(Bezdek1987)



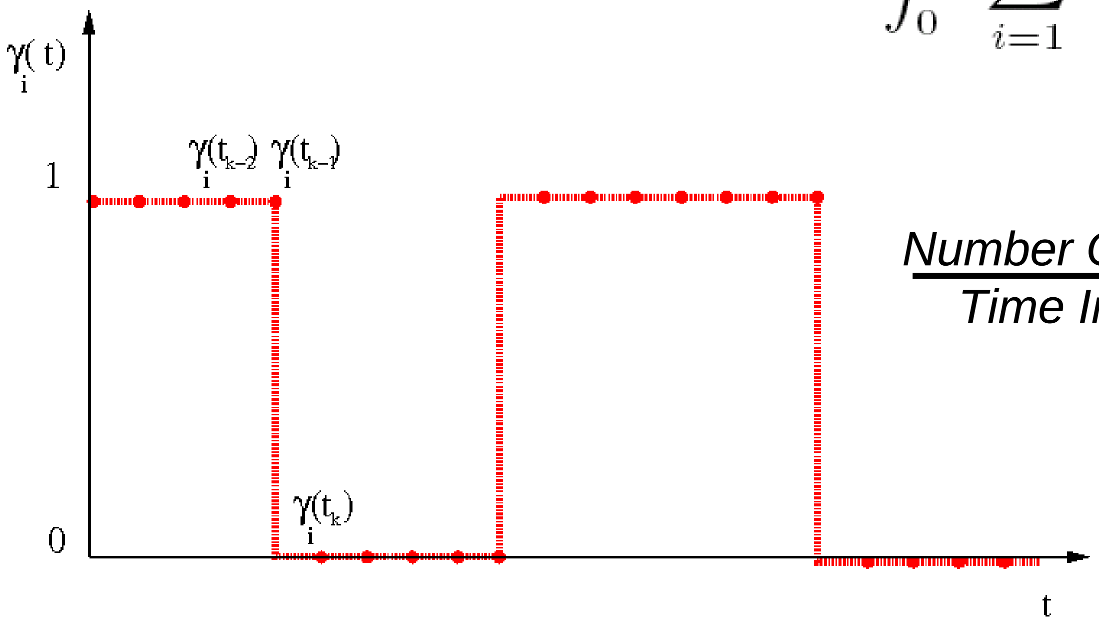
Just to "fuzzify" is not enough!

Problem 2: *identification of "persistent states"*

Let $\gamma_i(\cdot)$ on $t \in [0, T], i = 1, \ldots, \mathbf{K}$ be differentiable and

$\partial_t \gamma_i \in \mathcal{L}_2(0, T)$, i. e. :

$$\mathbf{L}(\Theta, \Gamma(t)) = \int_0^T \sum_{i=1}^{\mathbf{K}} \gamma_i(t) g(x_t, \theta_i) \to \min_{\Gamma(t), \Theta},$$



$$\frac{\textit{Number Of Jumps}}{\textit{Time Interval}} =$$

$$= \sum_{k=1} \frac{(\gamma_i(t_{k+1}) - \gamma_i(t_k))^2}{\Delta t}$$

Tikhonov-type Regularization

Problem 2: ***incorporation of temporal information***

Let $\gamma_i(\cdot)$ on $t \in [0,T], i = 1,\ldots,\mathbf{K}$ be differentiable and

$\partial_t \gamma_i \in \mathcal{L}_2(0,T)$, i. e. :

$$\mathbf{L}(\Theta,\Gamma(t)) = \int_0^T \sum_{i=1}^{\mathbf{K}} \gamma_i(t) g(x_t,\theta_i) \rightarrow \min_{\Gamma(t),\Theta},$$

subjected to

$$|\gamma_i|_{\mathcal{H}^1(0,T)} = \| \partial_t \gamma_i(\cdot) \|_{\mathcal{L}_2(0,T)} = \int_0^T (\partial_t \gamma_i(t))^2 \, dt \leq C_\epsilon^i < +\infty,$$

Regularized clustering functional:

$$\mathbf{L}^\epsilon(\Theta,\Gamma(t),\epsilon^2) = \mathbf{L}(\Theta,\Gamma(t)) + \epsilon^2 \sum_{i=1}^{\mathbf{K}} \int_0^T (\partial_t \gamma_i(t))^2 \, dt \rightarrow \min_{\Gamma(t),\Theta}$$

(H. 08, to appear in SISC)

FEM: Regularized Clustering Functional

Regularized clustering functional: *(H. 08, to appear in SISC)*

$$\mathbf{L}^{\epsilon}(\Theta, \Gamma(t), \epsilon^2) \quad = \quad \mathbf{L}(\Theta, \Gamma(t)) + \epsilon^2 \sum_{i=1}^{\mathbf{K}} \int_0^T \left(\partial_t \gamma_i\,(t)\right)^2 dt \longrightarrow \min_{\Gamma(t),\Theta}$$

FEM: Regularized Clustering Functional

Regularized clustering functional: (H. 08, to appear in SISC)

$$\mathbf{L}^{\epsilon}(\Theta,\Gamma(t),\epsilon^2) = \mathbf{L}(\Theta,\Gamma(t)) + \epsilon^2 \sum_{i=1}^{\mathbf{K}} \int_0^T \left(\partial_t \gamma_i(t)\right)^2 dt \rightarrow \min_{\Gamma(t),\Theta}$$

Galerkin-Ansatz :

$$\begin{aligned}\gamma_i(t) &= \tilde{\gamma}_i(t) + \delta_N \\ &= \sum_{k=1}^{N} \tilde{\gamma}_i^{(k)} v_k(t) + \delta_N\end{aligned}$$

where $\tilde{\gamma}_i^{(k)} = \int_0^T \gamma_i(t)\, v_k(t)\, dt$, and

$$\mathbf{L}^{\epsilon} = \tilde{\mathbf{L}}^{\epsilon} + \mathcal{O}(\delta_N) \rightarrow \min_{\tilde{\gamma}_i(t),\Theta},$$

$$\tilde{\mathbf{L}}^{\epsilon} = \sum_{i=1}^{\mathbf{K}} \int_0^T \left[\tilde{\gamma}_i(t)\, g(x_t,\theta_i) + \epsilon^2 \left(\partial_t \tilde{\gamma}_i(t)\right)^2\right] dt$$

FEM-Discretized Clustering Functional

(H. 08, to appear in SISC)

$$\tilde{\mathbf{L}}^{\epsilon} = \sum_{i=1}^{\mathbf{K}} \left[a^{\mathbf{T}}(\theta_i)\bar{\gamma}_i + \epsilon^2 \bar{\gamma}_i^{\mathbf{T}} \mathbf{H} \bar{\gamma}_i \right] \rightarrow \min_{\bar{\gamma}_i, \Theta}$$

subjected to

$$\sum_{i=1}^{\mathbf{K}} \tilde{\gamma}_i^{(k+1)} = 1, \quad \forall k = 1, \ldots, N,$$

$$\tilde{\gamma}_i^{(k+1)} \geq 0, \quad \forall k = 1, \ldots, N; i = 1, \ldots, \mathbf{K}.$$

Iterative Subspace Minimization: ***sparse QP*** can be used

where $a(\theta_i) = \left(\int_{t_1}^{t_2} v_1(t) g(x_t, \theta_i) dt, \ldots, \int_{t_{N-1}}^{t_N} v_N(t) g(x_t, \theta_i) dt \right)$

is a vector of *FEM-discretized model distances* and $\mathbf{H}$ is

a *mass-matrix* of the *FEM-basis*

Algorithm: monotony conditions

Theorem *Let for a given observed time series* $x(t) : \mathbf{R}^1 \rightarrow \Psi \subset \mathbf{R}^n$, *the model distance functional is chosen such that it satisfies (2),* Ψ *and* Ω *are compact,* $g(x_t, \cdot)$ *is continuously differentiable function of* θ *and*

$$\frac{\partial}{\partial \Theta} \tilde{\mathbf{L}}^{\epsilon} \left(\Theta^*, \bar{\gamma} \right) \quad = \quad 0,$$

has a unique solution $\Theta^* = \left(\theta_1^*, \ldots, \theta_{\mathbf{K}}^* \right), \theta_i * \in \Omega$ *for any fixed* $\bar{\gamma}$ *satisfying (18-19) and* $\frac{\partial^2}{\partial \Theta^2} \tilde{\mathbf{L}}^{\epsilon} \left(\Theta^*, \bar{\gamma} \right)$ *is positive definite. Then for any* $\epsilon^2 \geq 0$ *and any finite continuous non-negative finite elements set* $\{v_1(t), v_2(t), \ldots, v_N(t)\} \in \mathcal{L}_2(0, T)$ *such that the respective* mass-matrix $\mathcal{H}$ *is positive definite, the above algorithm is* monotonous, *i. e., for any* $j \geq 1$

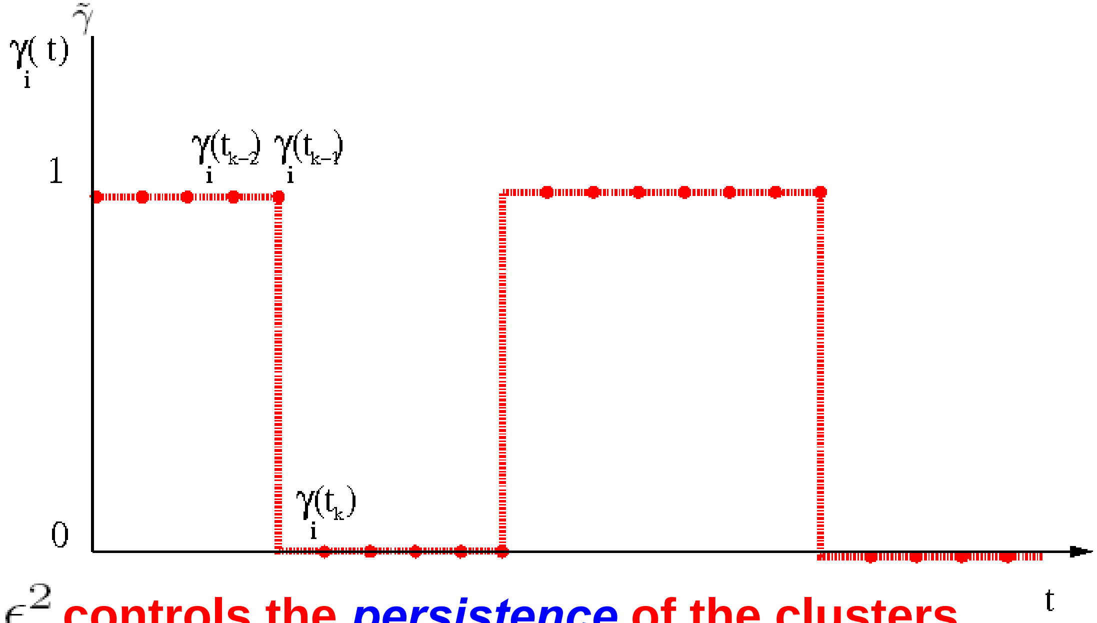
$$\tilde{\mathbf{L}}^{\epsilon} \left(\Theta^{[j+1]}, \bar{\gamma}_i^{[j+1]} \right) \quad \leq \quad \tilde{\mathbf{L}}^{\epsilon} \left(\Theta^{[j]}, \bar{\gamma}_i^{[j]} \right).$$

Convergence to a local optimum only!

(coupling to some *global opimizer* necessary)

Regularity and Persistence

$\tilde{\gamma}^{\epsilon} = \arg\min \tilde{\mathbf{L}}^{\epsilon}(\Theta, \tilde{\gamma})$ ⟷ *Exit Times* τ_i^{exit}



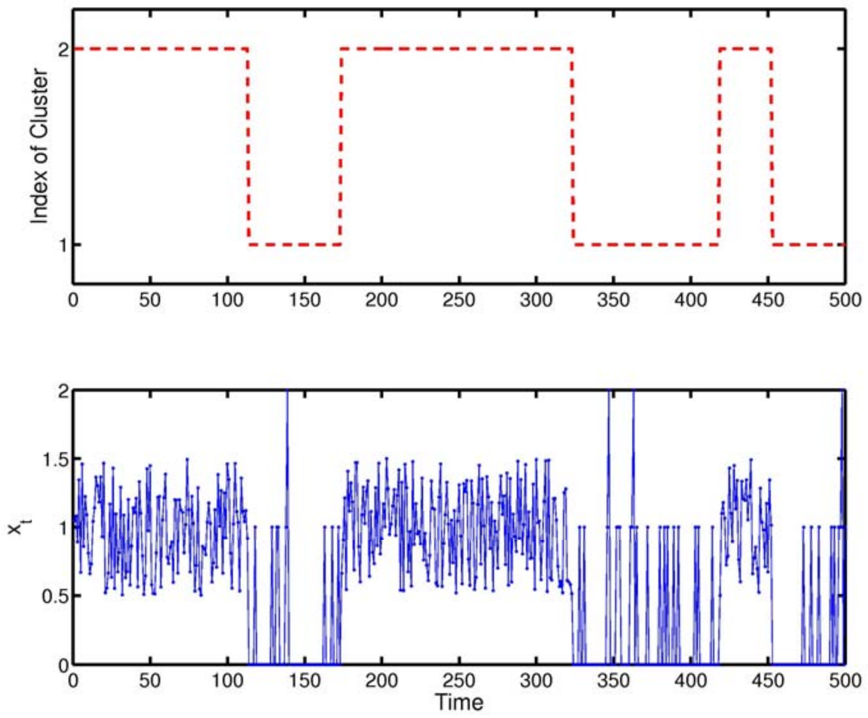
ϵ^2 **controls the *persistence* of the clusters**

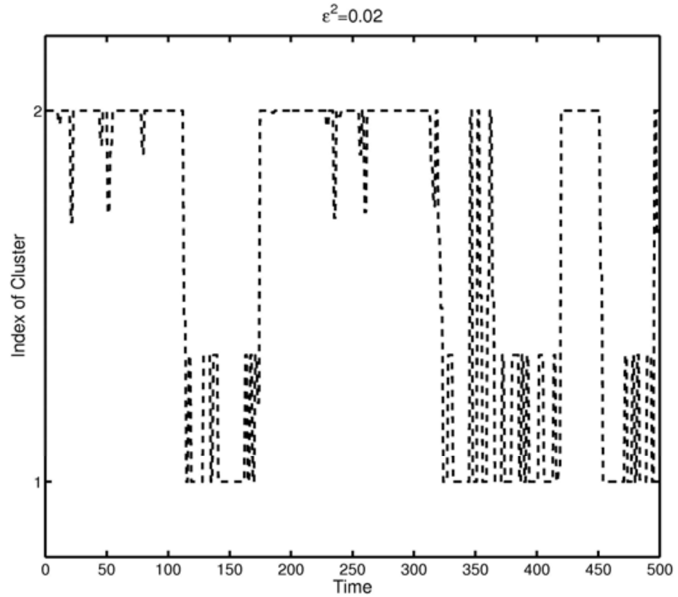
<u>Exercise 2</u>*: calculate the mean time the Markov chain with a given stationary transition probability matrix P will spend in the state i.*

Toy Example I

$$\mathbf{L}\left(\gamma_i(t), \mathbf{T}_i, \mu_i\right) \rightarrow \min$$

$$g\left(x, \theta_i\right) \quad = \quad \| x - \theta_i \|^2,$$

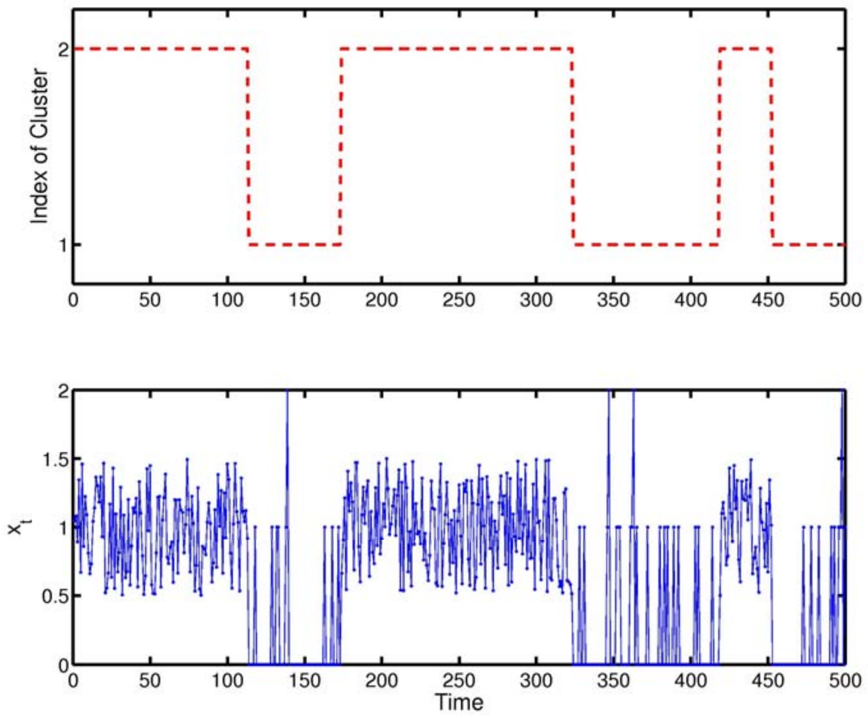


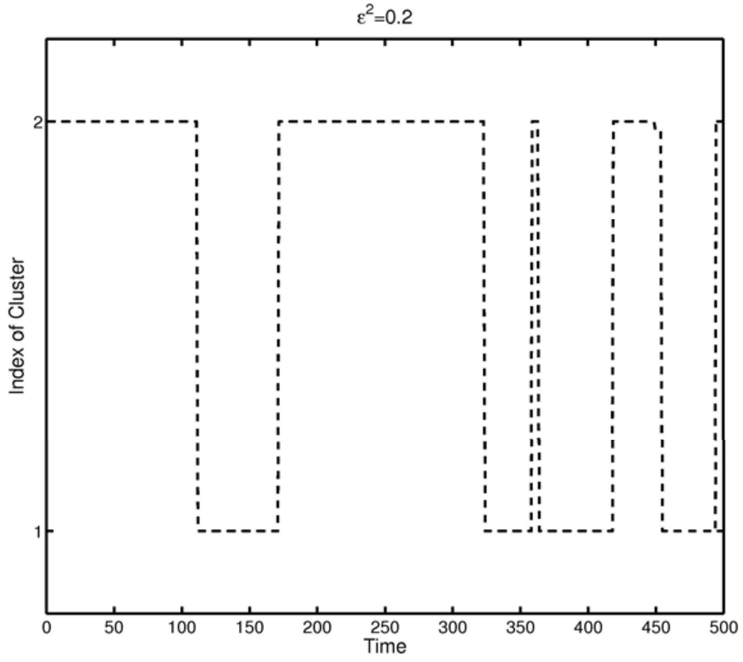


Toy Example I

$$\mathbf{L}\left(\gamma_i(t), \mathbf{T}_i, \mu_i\right) \rightarrow \min$$

$$g\left(x, \theta_i\right) \quad = \quad \| \, x - \theta_i \, \|^2,$$

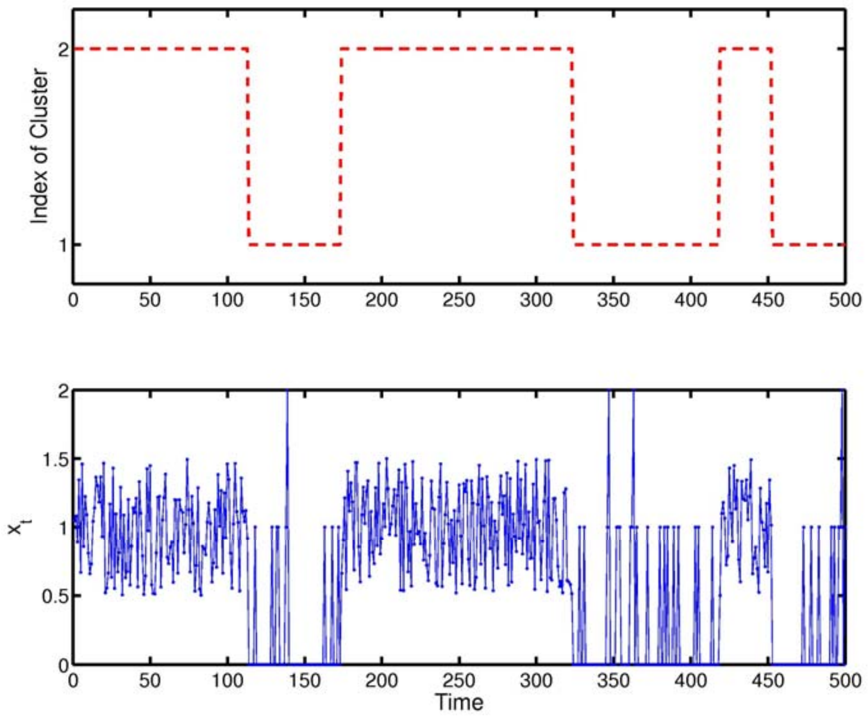


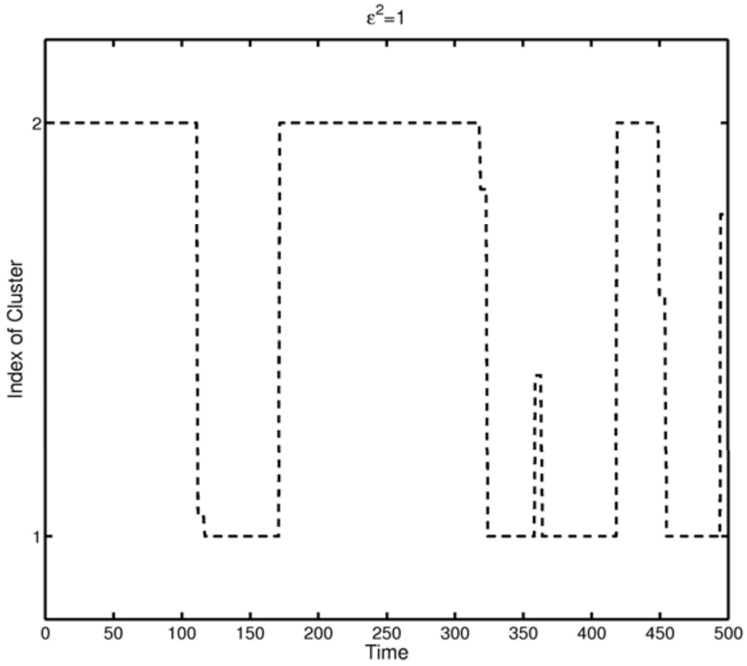


Toy Example I

$$\mathbf{L}\left(\gamma_i(t), \mathbf{T}_i, \mu_i\right) \longrightarrow \min$$

$$g\left(x, \theta_i\right) \quad = \quad \| \, x - \theta_i \, \|^2,$$

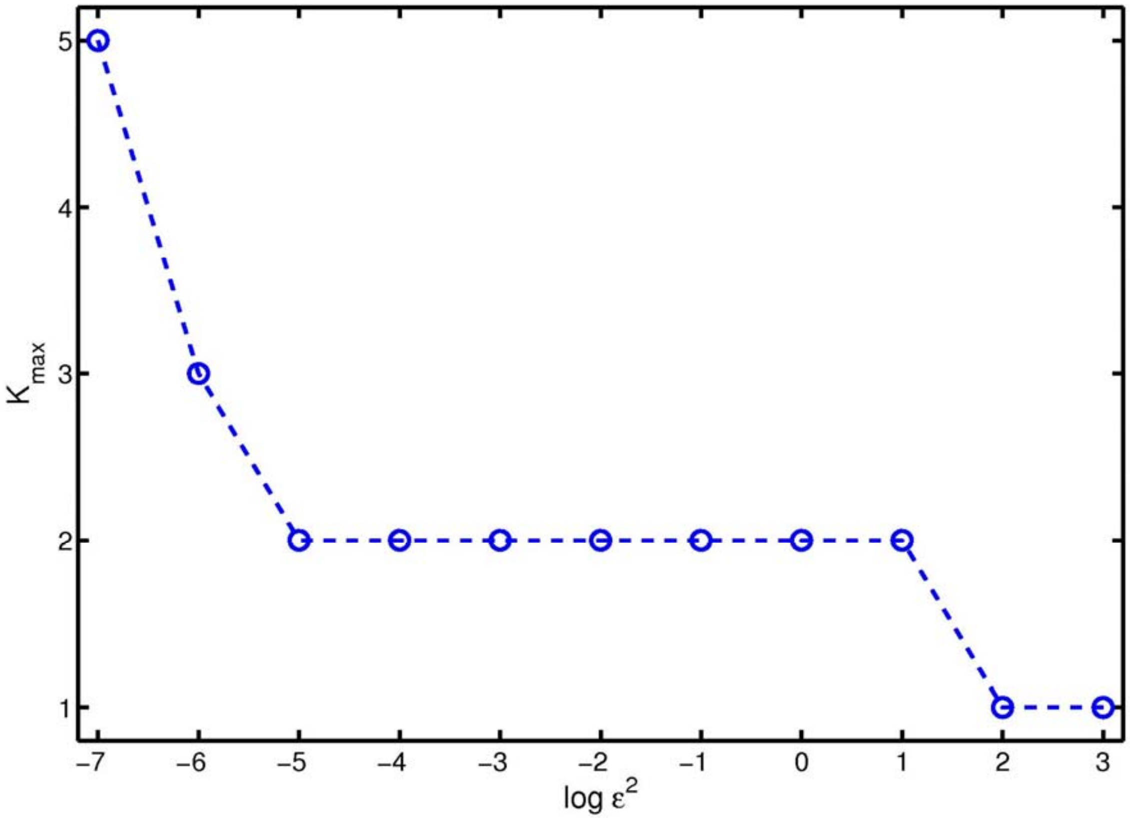




Toy Example I

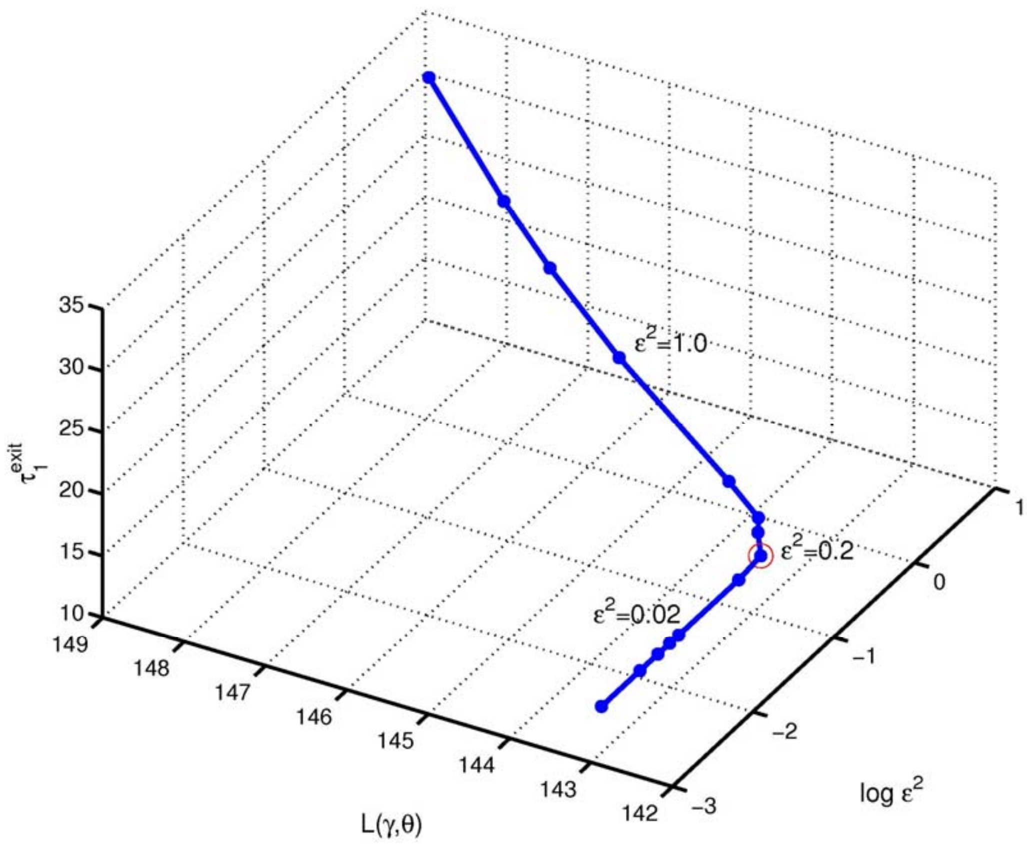
How to determine the optimal K: probabilistic model assumptions a posteriory

$$g(x, \theta_i) = \| x - \theta_i \|^2,$$



Toy Example I

How to determine the optimal ε: standard L-Curve approach from Tikhonov-regularized linear least-squares problems (Cullum(79), Hansen(99))

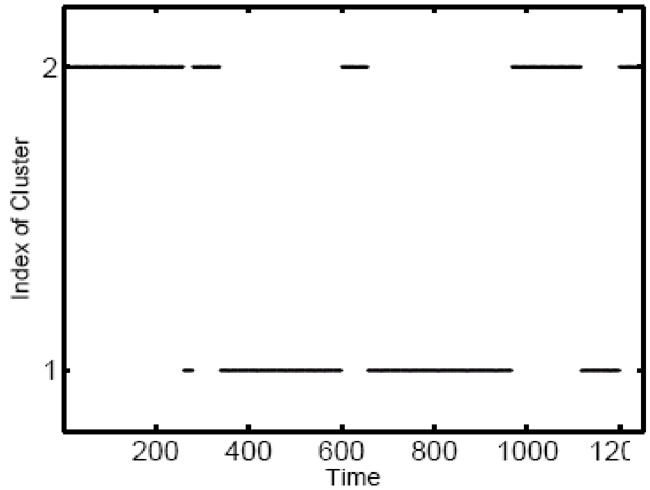


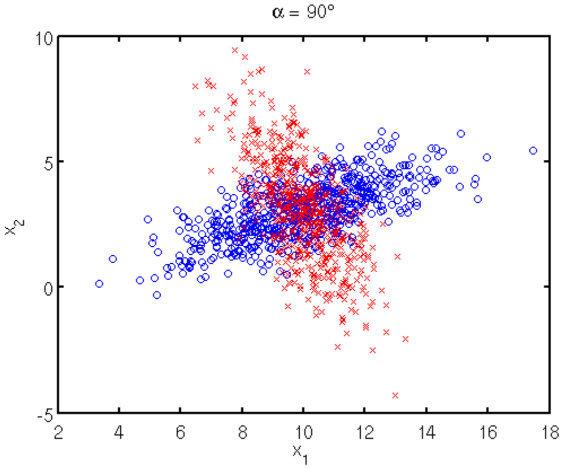
Toy Example II

$$\mathbf{L}\left(\gamma_i(t), \mathbf{T}_i, \mu_i\right) \to \min$$

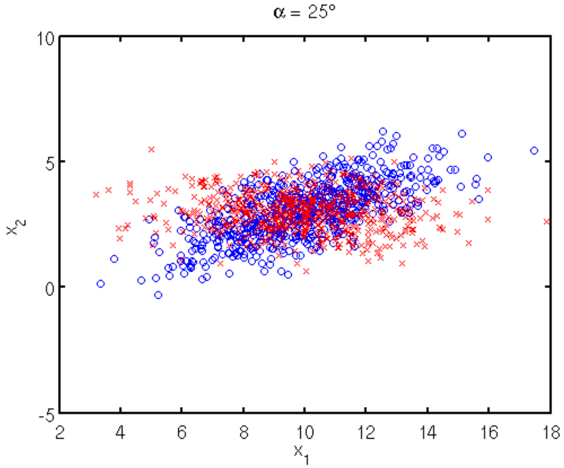
$$g\left(x, \theta_i\right) \quad = \quad \| \, x - \mathcal{T}_i \mathcal{T}_i^{\mathbf{T}} x \, \|^2$$

Switching between the distributions



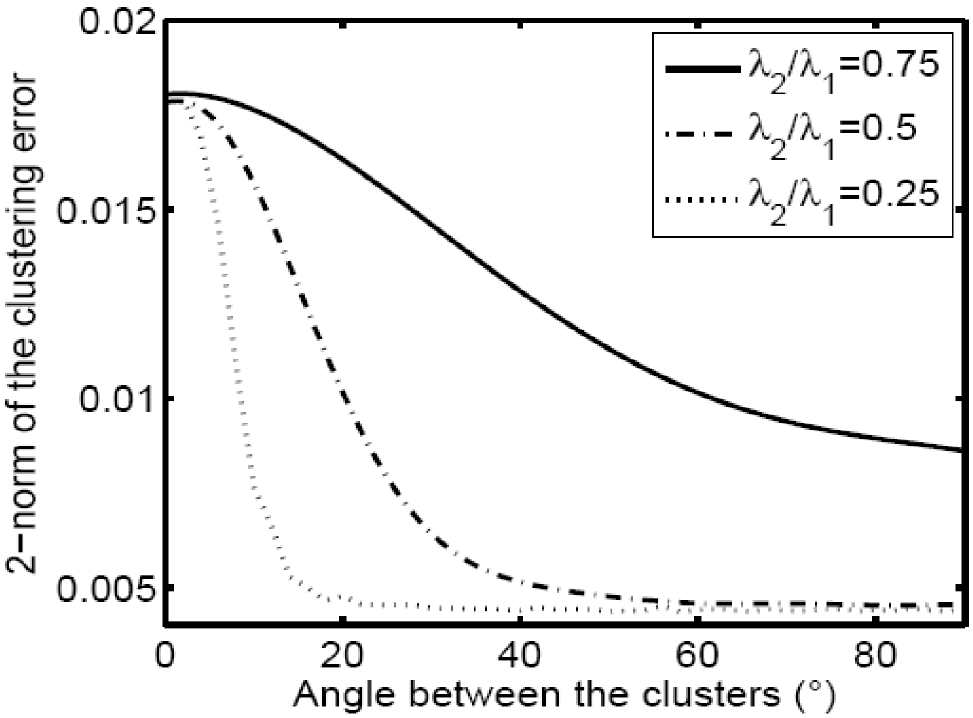


Angle α between the distributions



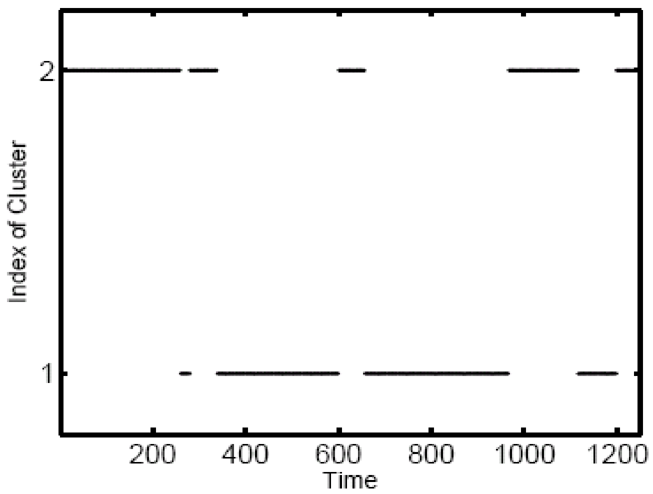
Toy Example

Error of FEM-Based Clustering

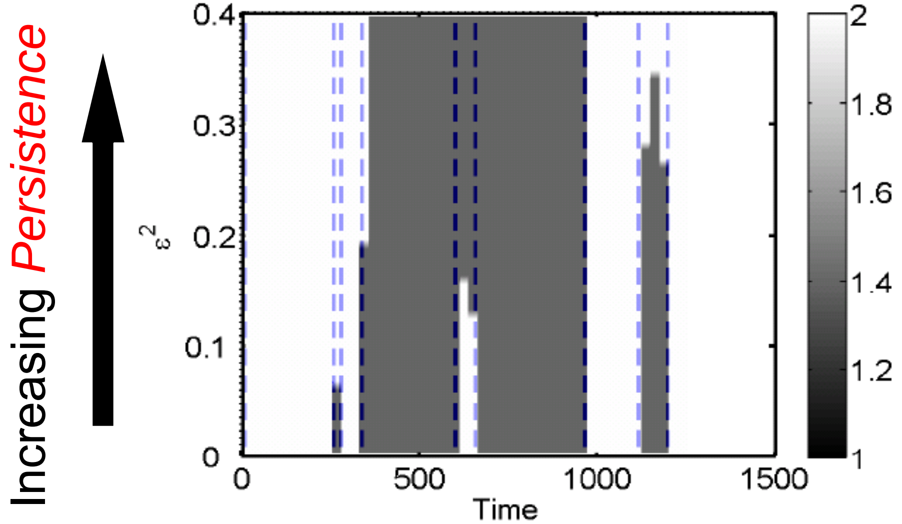


Toy Example

Effect of Regularization



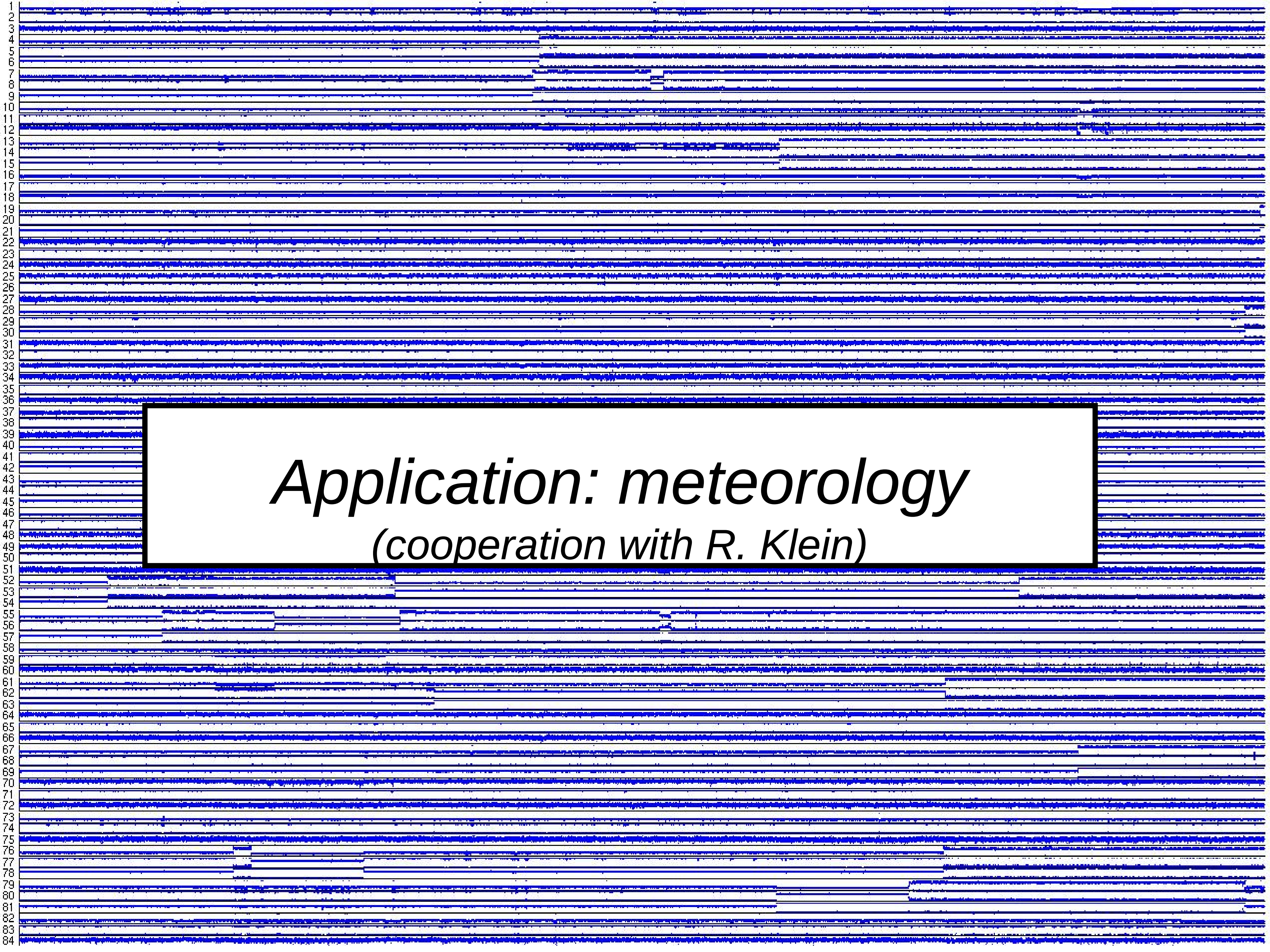
Original switching between the distributions



Switching between the distributions identified by

$$\mathbf{L}\left(\gamma_i(t), \mathbf{T}_i, \mu_i\right) \longrightarrow \min$$

$$g\left(x, \theta_i\right) \quad = \quad \| \, x - \mathcal{T}_i \mathcal{T}_i^{\mathbf{T}} x \, \|^2$$

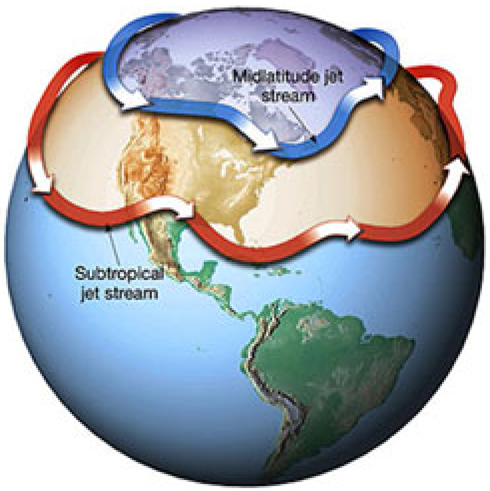


Application: meteorology

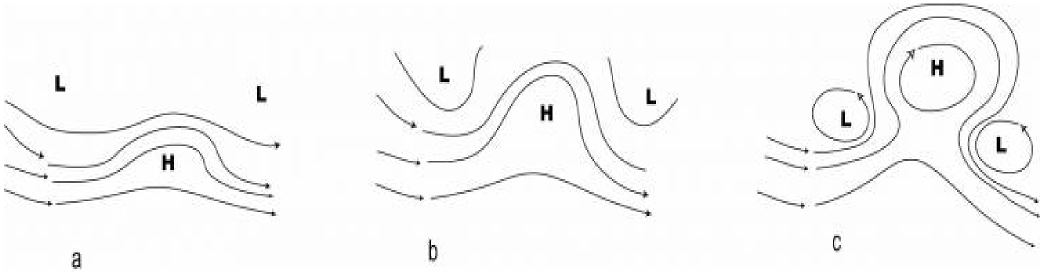
(cooperation with R. Klein)

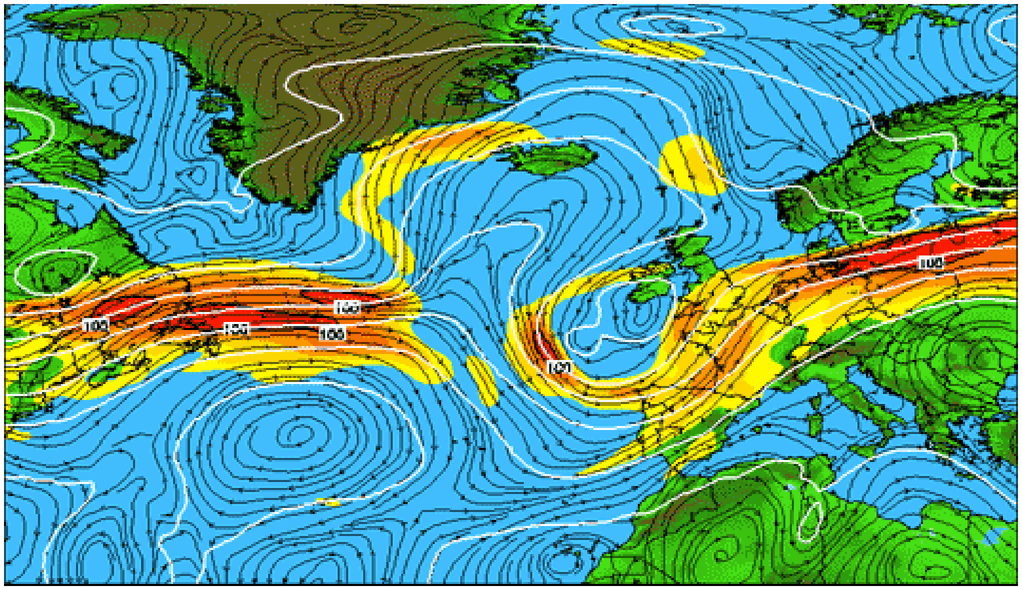
“Hidden“ Weather Phases: motivation

Wind Jets transport moisture from US to Europe



Jet-Blocking





Jet-Blocking results in “indian summer“ and “blackberry cold“ in Europe

$$\mathbf{L}\left(\gamma_i(t), \mathbf{T}_i, \mu_i\right) \to \min$$: K=4

SPP 1276 "MetStröm" Project

Weather Data in Europe: 29x20 grid (44 years)

(Data from H.Osterle, PIK)

Up to 4 hidden states are statistically separable

Metastable | **Intermediate** | **Metastable**

Looking through Markovian  **: analysis of exit times**

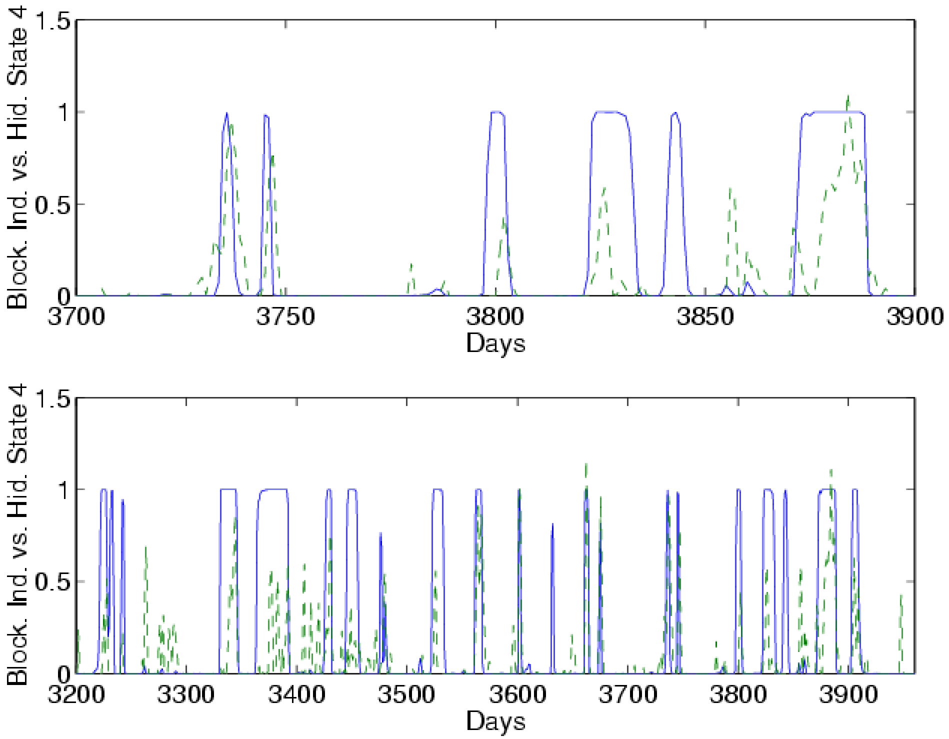
$$\mathbf{L}\left(\gamma_i(t), \mathbf{T}_i, \mu_i\right) \rightarrow \min \; : K=4$$

Weather Data in Europe: 29x20 grid (44 years)

(Data from H.Osterle, PIK)

Up to 4 hidden states are statistically separable

SPP 1276 "MetStröm" Project



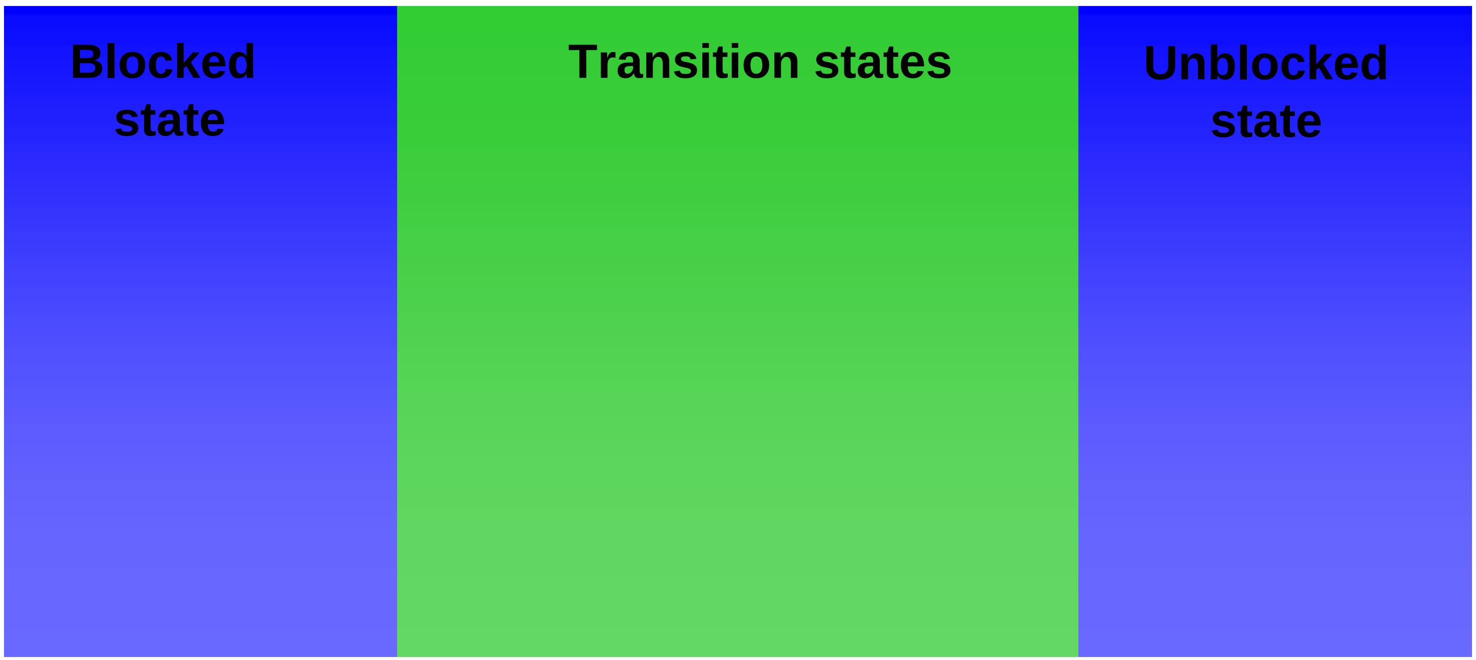
Comparison with *Lejenas-Okland blocking index*

Hidden State 4: ***Jet Blocking Situation***

$$\mathbf{L}\left(\gamma_i(t), \mathbf{T}_i, \mu_i\right) \rightarrow \min$$: K=4

Stochastic Prediction: Markov+SDEs (*H./Klein/Dolaptchiev/Schütte, SIAM MMS 06,*
H., JAS 08,
H./Dolaptchiev/Eliseev/Mokhov/Klein, JAS 08)

SPP 1276 "MetStröm" Project



Reduced dynamical description

$$z(t) = \mathbf{T}_i^{\top}\left(x_t - \mu_i\right)$$
$$\dot{z}(t) \quad = \quad F_i\left(z(t) - \bar{\mu}_i\right) + \Sigma_i \dot{W}(t)$$

veritas iustitia libertas

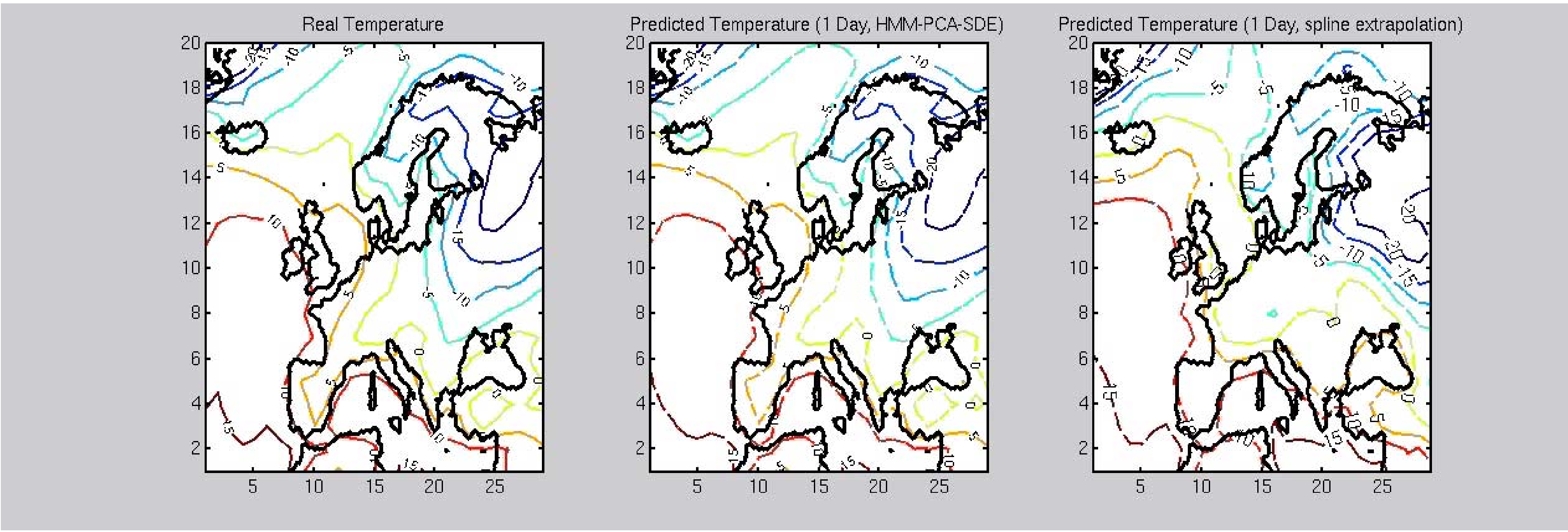
$\mathbf{L}(\gamma_i(t), \mathbf{T}_i, \mu_i) \rightarrow \min$ + SDE Prediction

Weather Data in Europe: 29x20 grid (44 years)

(Data from H.Osterle, PIK)

Up to 4 hidden states are statistically separable

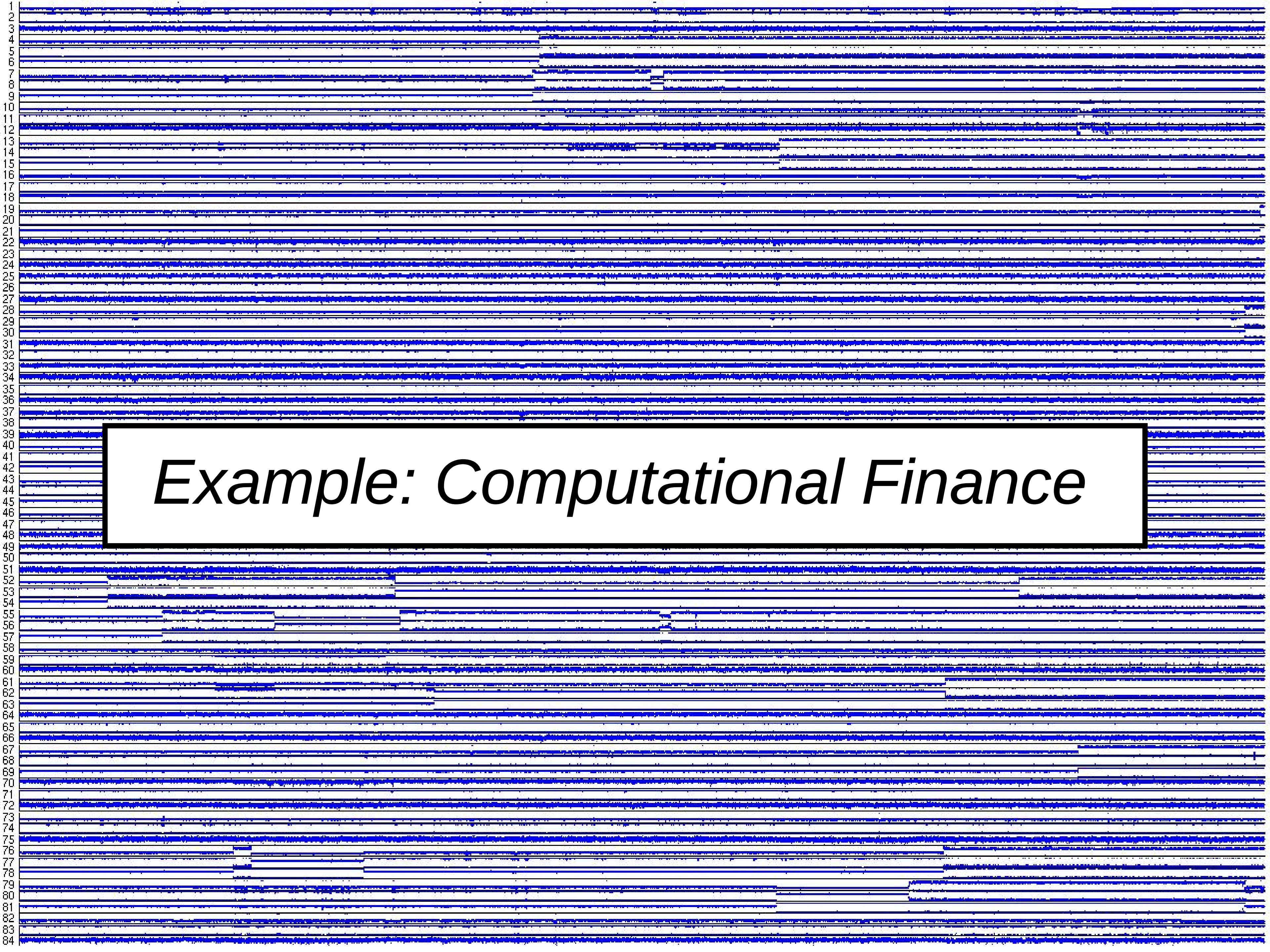
1 Day temperature predictions



$$z(t) = \mathbf{T}_i^\top (x_t - \mu_i)$$

$$\dot{z}(t) = F_i (z(t) - \bar{\mu}_i) + \Sigma_i \dot{W}(t)$$

Example: Computational Finance

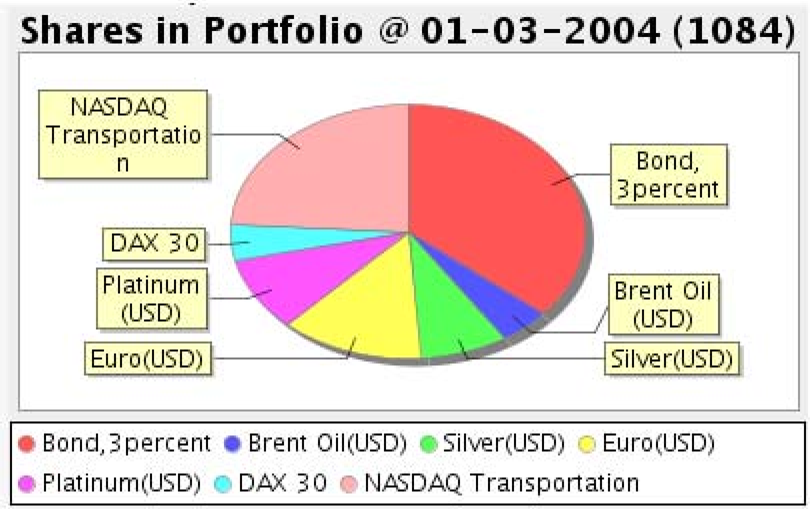


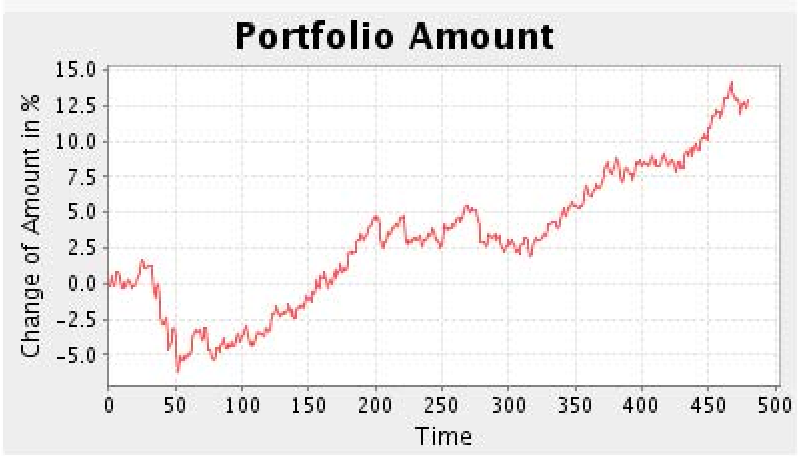
Minimal Risk Portfolios

Optimal Portfolio: Maximize Yield and Minimize Risk

(H. Markowitz 1952, Nobel Price 1990)

Mathematical Analysis of available Financial Information





Memo II: Stochastic Process

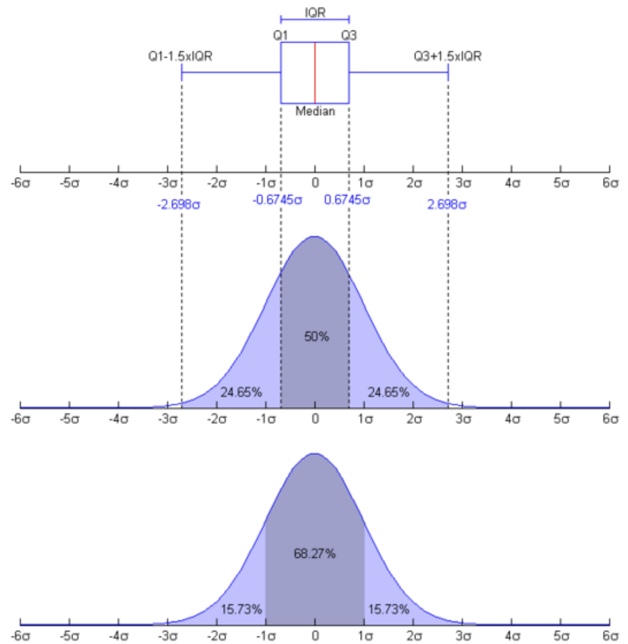
Given a probability space $(\Omega, \mathcal{F}, P)$, a **stochastic process** (or **random process**) with state space X is a collection of X-valued random variables indexed by a set T ("time"). That is, a stochastic process F is a collection

$$\{F_t : t \in T\}$$

where each F_t is an X-valued random variable.

Probability Density Function:

$$\Pr(4.3 \le X \le 7.8) = \int_{4.3}^{7.8} f(x)\, dx$$



Expectation Value:

$$\mathrm{E}(X) = \int_\Omega X \, \mathrm{d}P \qquad \mu = \int x\, p(x)\, dx$$

Variance:

$$\mathrm{Var}(X) = \int (x-\mu)^2\, p(x)\, dx$$

Application: Minimal Risk Portfolio

Let $P_s(T) \in \mathbf{R}^n$ be a *stochastic price process* and $x \in \mathbf{R}^n$

is a *portfolio*, then $c(T,x) = \langle P_s(T), x \rangle$ the *portfolio price*

Maximize Yield and ***Minimize Risk***

Application: Minimal Risk Portfolio

Let $P_s(T) \in \mathbf{R}^n$ be a *stochastic price process* and $x \in \mathbf{R}^n$

is a *portfolio*, then $c(T,x) = \langle P_s(T), x \rangle$ the *portfolio price*

$$\mathbf{E}(c) - \alpha \mathbf{V}ar^{-}(c - \mathbf{E}(c)) \quad \rightarrow \quad \max_x$$

Application: Minimal Risk Portfolio

Let $P_s(T) \in \mathbf{R}^n$ be a *stochastic price process* and $x \in \mathbf{R}^n$

is a *portfolio*, then $c(T,x) = \langle P_s(T), x \rangle$ the *portfolio price*

$$\mathbf{E}(c) - \alpha \mathbf{V}ar^{-}(c - \mathbf{E}(c)) \quad \to \quad \max_x$$

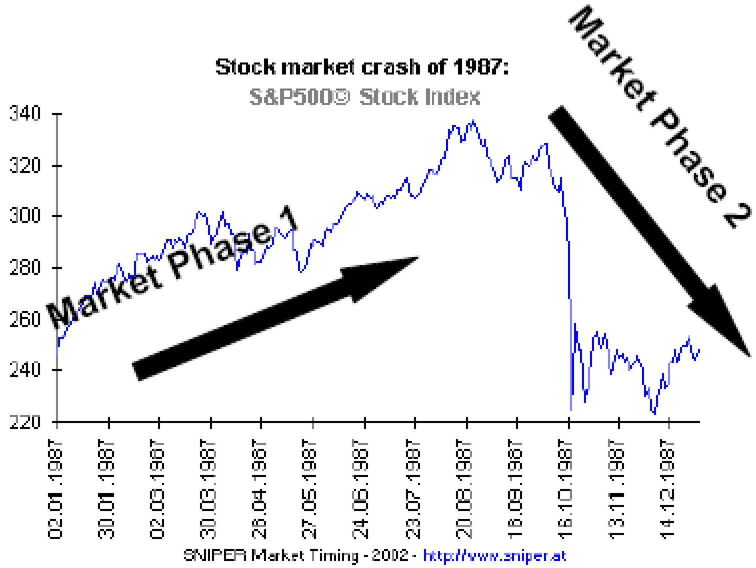
Numerical solution possible

Exercise 3: *write the above maximization problem in vector-matrix form. What kind of constraints on x will be meaningful? Under which conditions on* $\partial P_s(t)/\partial t$ *does this constrained maximization problem have a unique solution for a fixed* α

Minimal Risk Portfolios

Persistent Regimes in Financial Data: Market Phases

Market Phases are hidden in *multiple*, *noisy* data-series

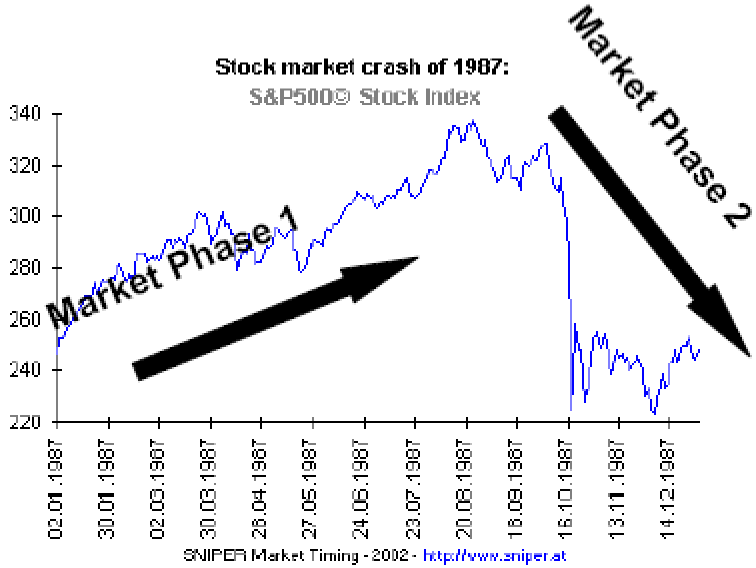


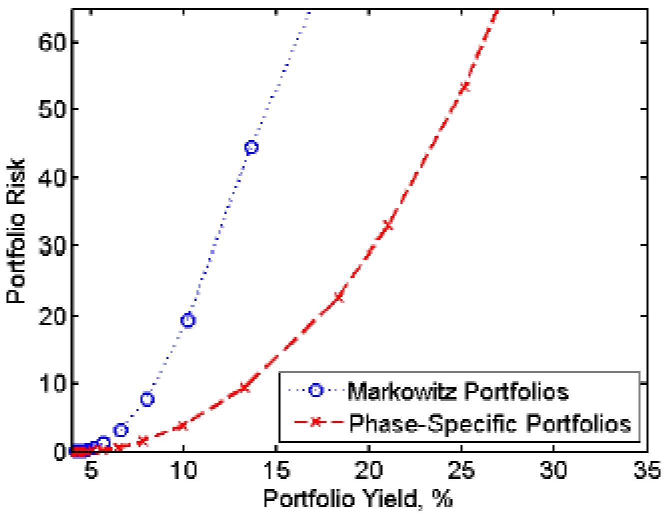
I. Horenko, Ch. Schütte,
German Patent 10 2007 014 921.4 on 22.03.2007

Minimal Risk Portfolios

Persistent Regimes in Financial Data: Market Phases

Market Phases are hidden in *multiple*, *noisy* data-series





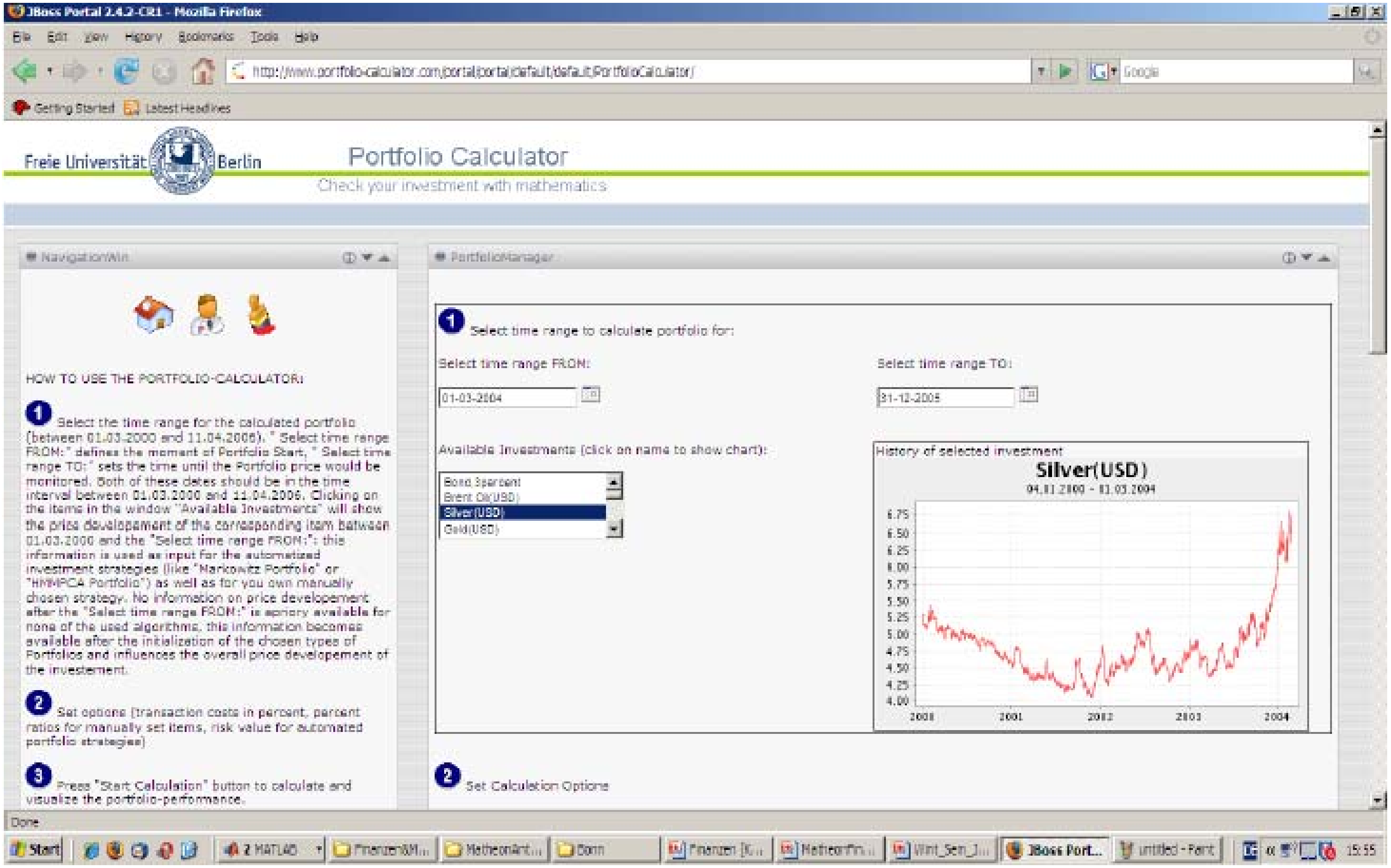
Identification of *hidden market phases* reduces the *portfolio risk*

$$\mathbf{L}\left(\gamma_i(t), \mathbf{T}_i, \mu_i\right) \longrightarrow \min$$

$$g\left(x, \theta_i\right) = \| x - \mathcal{T}_i \mathcal{T}_i^{\mathbf{T}} x \|^2$$

I. Horenko, Ch. Schütte,
German Patent 10 2007 014 921.4 on 22.03.2007

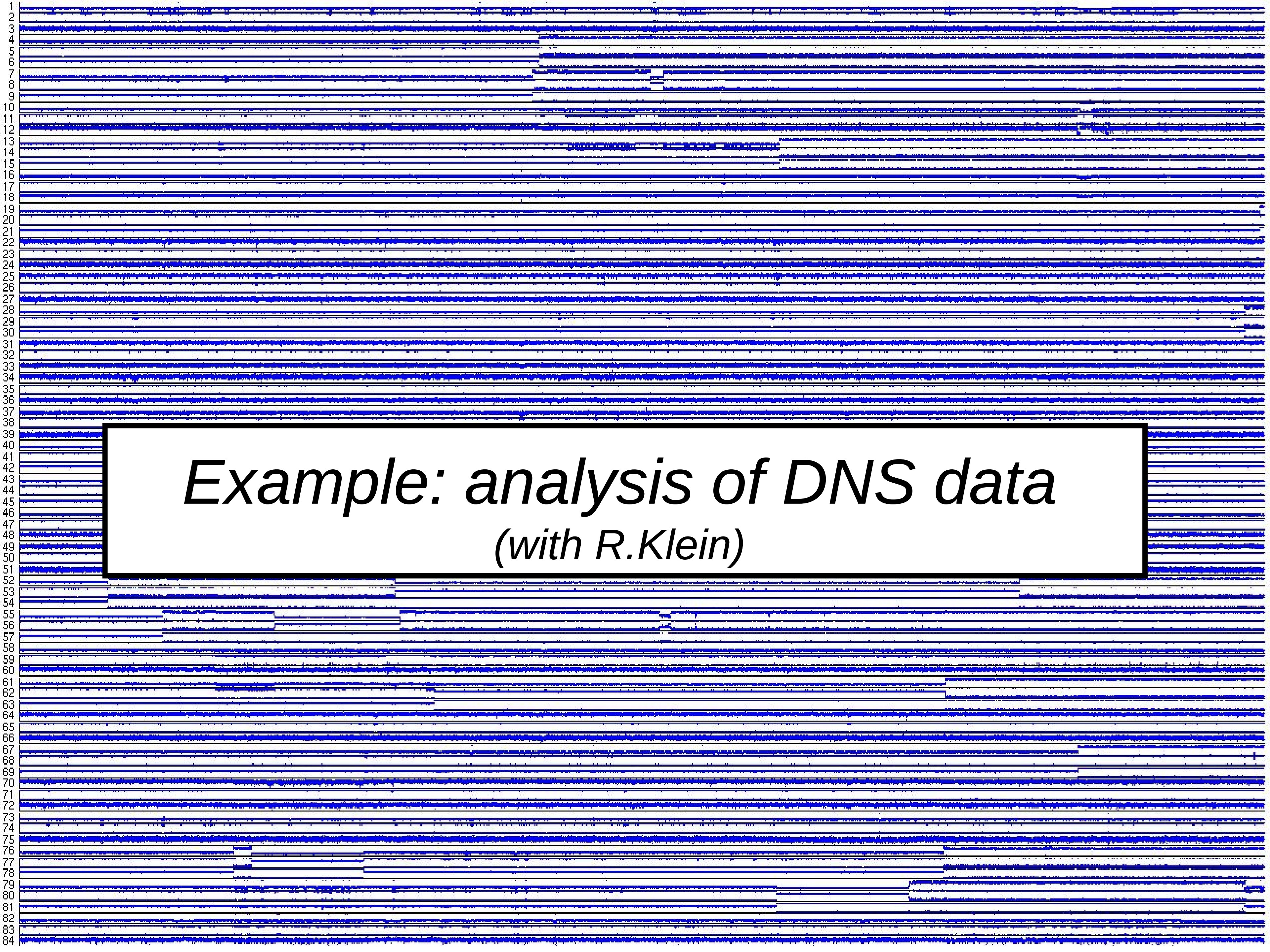
http://www.portfolio-calculator.com



http://www.portfolio-calculator.com

$$\mathbf{L}\left(\gamma_i(t), \mathbf{T}_i, \mu_i\right) \rightarrow \min$$

$$g\left(x, \theta_i\right) \quad = \quad \| \, x - \mathcal{T}_i \mathcal{T}_i^{\mathbf{T}} x \, \|^2$$



Example: analysis of DNS data

(with R.Klein)

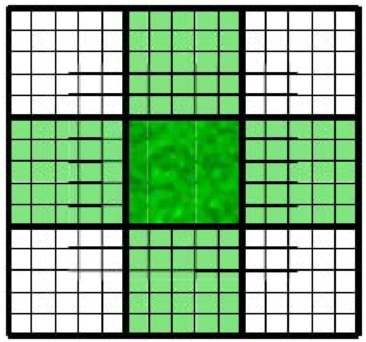
Discrete-Continuous Hybrid Models…

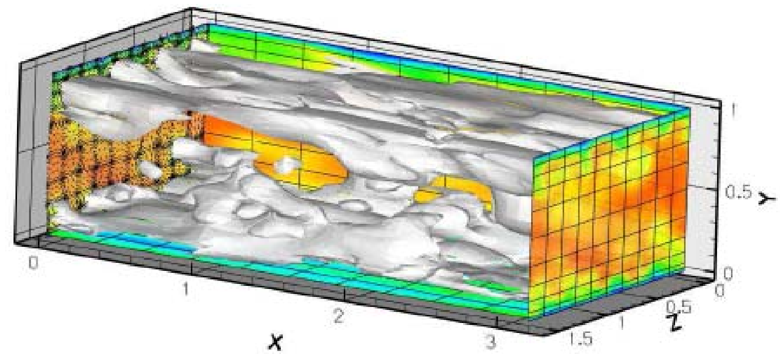
SPP 1276 "MetStröm" Project

fine-grid spatial patterns ⇔ coarse-grid stencil data

Issues (1st phase):

- Structure of fluctuations
- Fluctuations vs. mean flow
- Coarse-grid dynamics





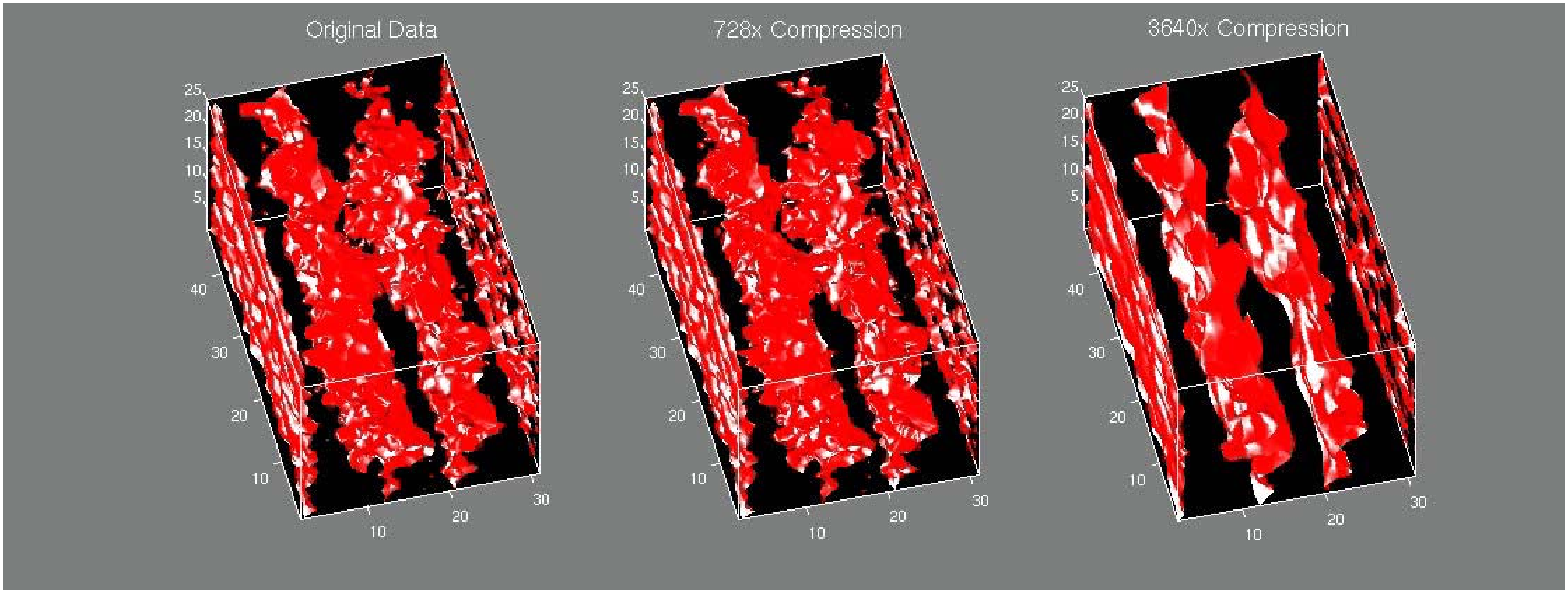
DNS by G. Gassner (IAG, Uni Stuttgart)

Aim: construction of *reduced stochastic models* for subscale phenomena

based on available DNS data *or data-based model reduction*

DNS Data: Compression

Fluid Flow in 3D
1.2 TB Data (Data from R.Klein/M.Uhlmann, FU)

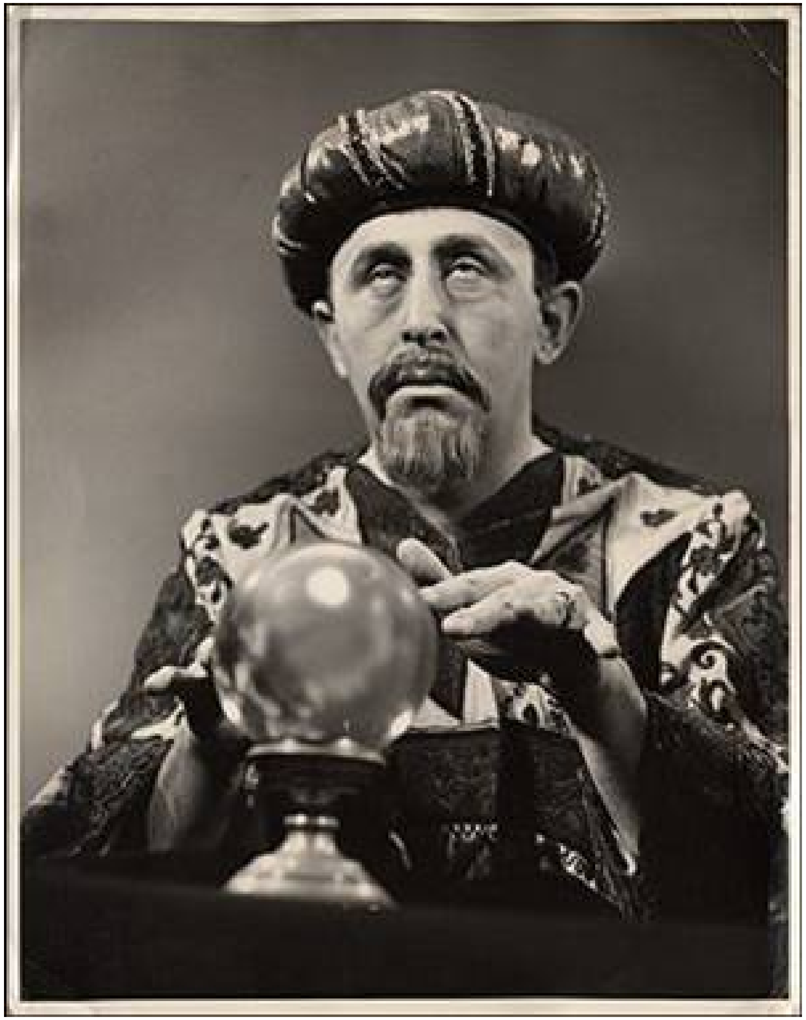


15 x 80 GB ≈ 2.000 EUR **1 USB with 2 GB ≈ 20 EUR** **1 CD ≈ 0.2 EUR**

$$z(t) = \mathbf{T}_i^{\top}\left(x_t - \mu_i\right)$$
$$\dot{z}(t) = F_i\left(z(t) - \bar{\mu}_i\right) + \Sigma_i \dot{W}(t)$$

Take–Home-Messages :

1. *Variational approach* to time series analysis. No explicit probabilistic assumptions needed (difference to GMM/HMM)
2. Probabilistic assumptions can be done a posteriory (determination of ***K***)
3. *Regularization* controls the *persistence* of regimes.



Thank you for attention!