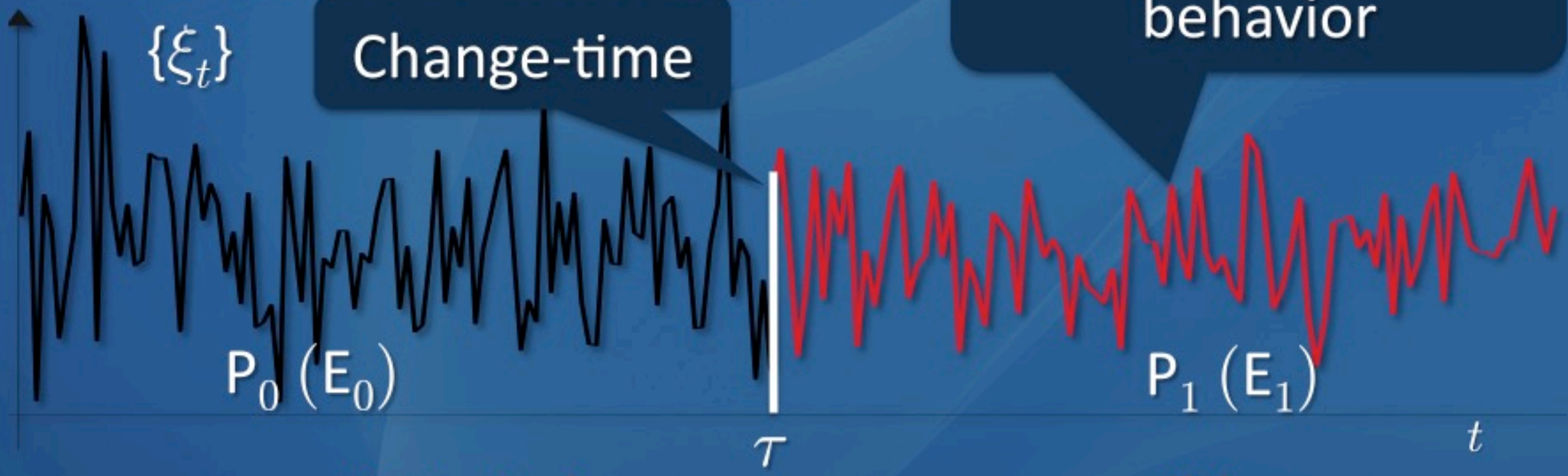
Metrics and Optimum Tests
for Sequential Change-Detection

George V. Moustakides
University of Patras, Greece
& Rutgers University, USA

Outline

- Definition of the change detection problem, Applications, Questions with existing formulations
- Model for mechanism that imposes change
- Two generic setups for sequential detection of changes
  - Complete knowledge
  - Incomplete knowledge

Problem definition



Detect change as soon as possible

Data become available sequentially: at each instant t obtain new sample ξ_t.

Detector: Every instant t consult available data ξ_1 ,..., ξ_t
Use them to make **Binary Decision**

Each instant t decide between:

- ~~A change took place before and including t.~~
- ~~A change didn't take place before and including t.~~

STOP

Ask for more observations

There will be a point in time (call it) T **where we stop**.
T is random time controlled by the observations.

Need a test to implement binary decision at each time t:

$$T = \inf\{G_t(\xi_1, \ldots, \xi_t) \geq \nu\}$$

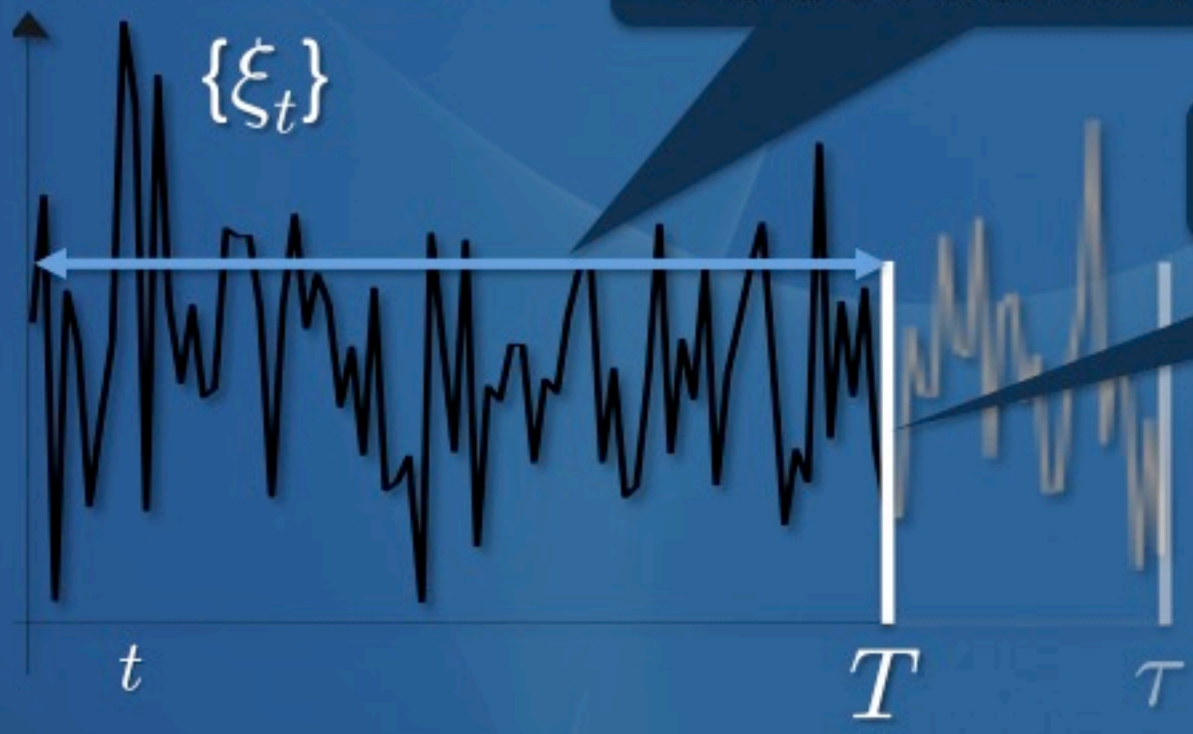
$G_t(\xi_1, \ldots, \xi_t)$ Stop $\gtrless$ Sample ν

Stopping rule

Stopping Time *adapted* to the observation history

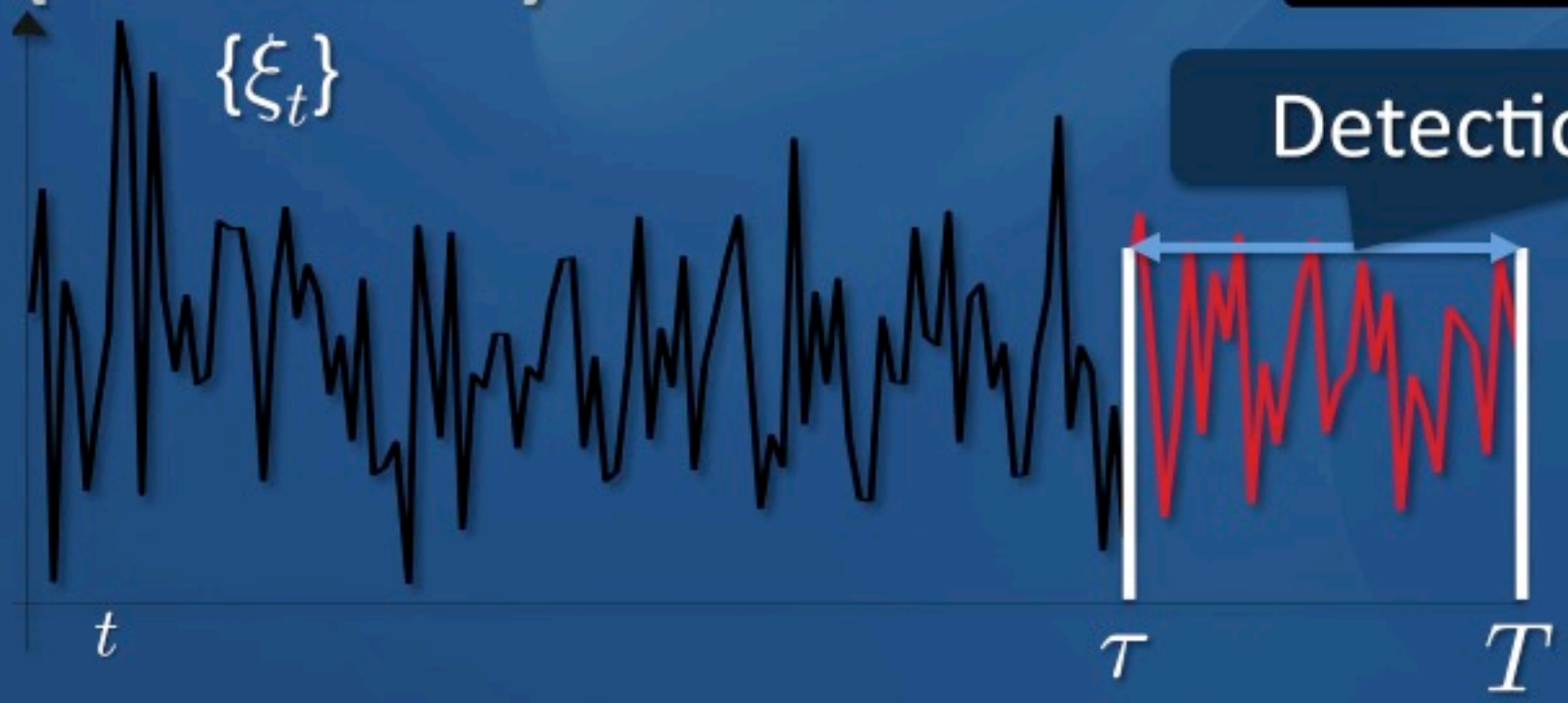
Class of stopping times: **Extremely rich**

False Alarm



Design T to **minimize** Detection Delay $(T - \tau)$ while **controlling** False Alarms

(Successful) Detection



Quality monitoring of manufacturing process

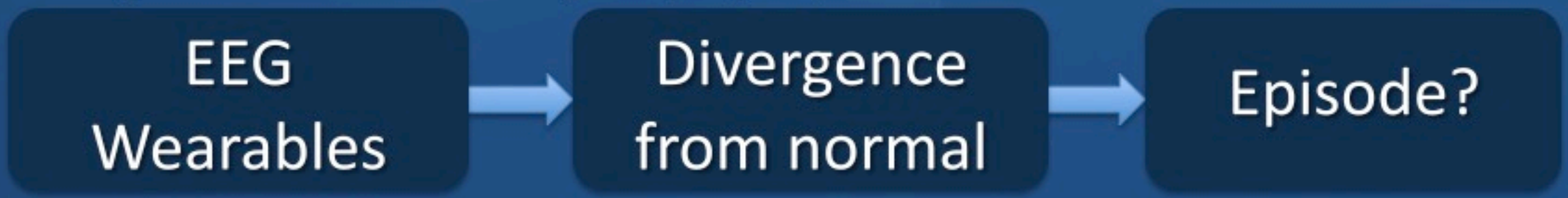


Medical Applications

Epidemic Detection



Early Detection of Epilepsy Episode



Metrics and Optimum Tests for Sequential Change-Detection, ISyE, GaTech, Dec. 2016

Financial Applications

- Structural Change-detection in Exchange Rates
- Portfolio Monitoring

Electronic Communications
Seismology
Speech & Image Processing (segmentation)
Vibration monitoring (Structural health monitoring)
Security monitoring (fraud detection)
Spectrum monitoring
Scene monitoring
Network monitoring (router failures, attack detection)
⋮

CUSUM: 3,000 hits in 2015. Google Scholar.
80% in Change Detection: 2300 articles

Metrics and Optimum Tests for Sequential Change-Detection, ISyE, GaTech, Dec. 2016

Performance Metrics

$$J(T) = \mathsf{E}_1[T - \tau \,|\, T > \tau]$$

What is the change-time τ ?

Metric must measure only success

Failures will be dealt through False Alarm control

Shiryaev (1963): τ is random with known prior:

$$J_{\mathsf{S}}(T) = \mathsf{E}_1[T - \tau \,|\, T > \tau]$$

Too restrictive!

Pollak (1985): τ is deterministic and unknown:

$$J_{\mathsf{P}}(T) = \sup_{t \geq 0} \mathsf{E}_1[T - t \,|\, T > t]$$

Leads to SR test

Lorden (1971): τ is deterministic and unknown:

$$J_{\mathsf{L}}(T) = \sup_{t \geq 0} \sup_{\xi_1, \ldots, \xi_t} \mathsf{E}_1[T - t \,|\, T > t, \xi_1, \ldots, \xi_t]$$

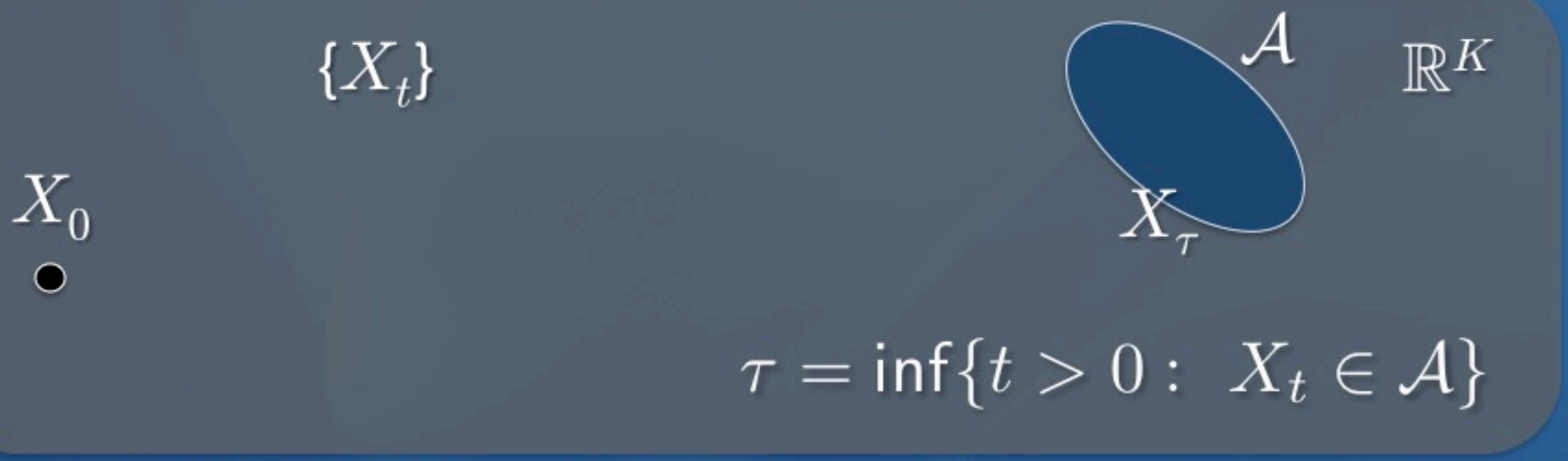
Leads to CUSUM

Too pessimistic (?)

Model for change imposing mechanism

A random vector process $\{X_t\}$ evolves in time in $\mathbb{R}^K$

$\mathcal{A}$ is a subset in $\mathbb{R}^K$



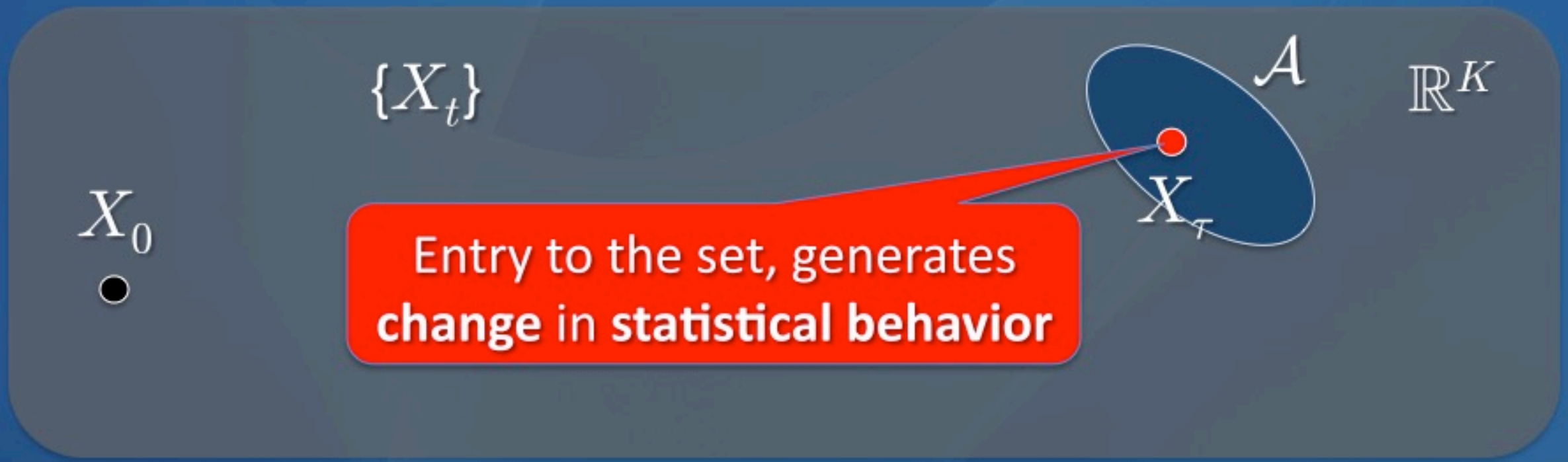
τ is a first entry time, depends on $\{X_t\}$.

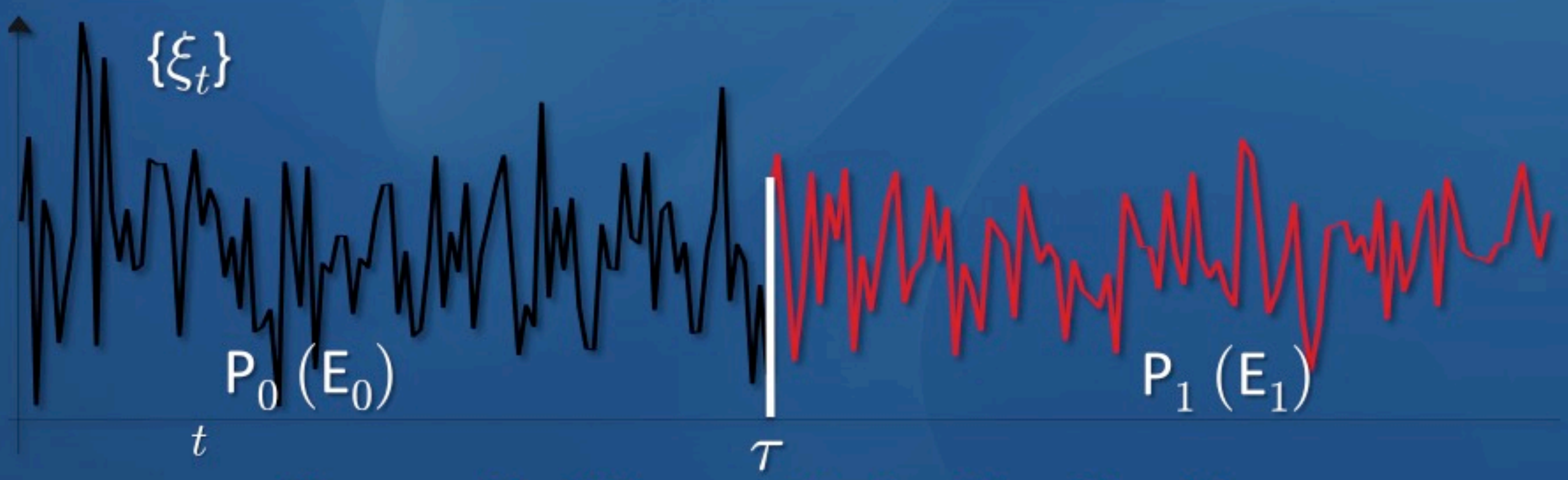
We would like to detect it.

If $\{X_t\}$ observable and $\mathcal{A}$ known, problem is trivial.

If $\{X_t\}$ (partially) hidden and/or $\mathcal{A}$ unknown, problem is challenging.

Instead of $\{X_t\}$ we observe process $\{\xi_t\}$





Sequential Change Detection Problem

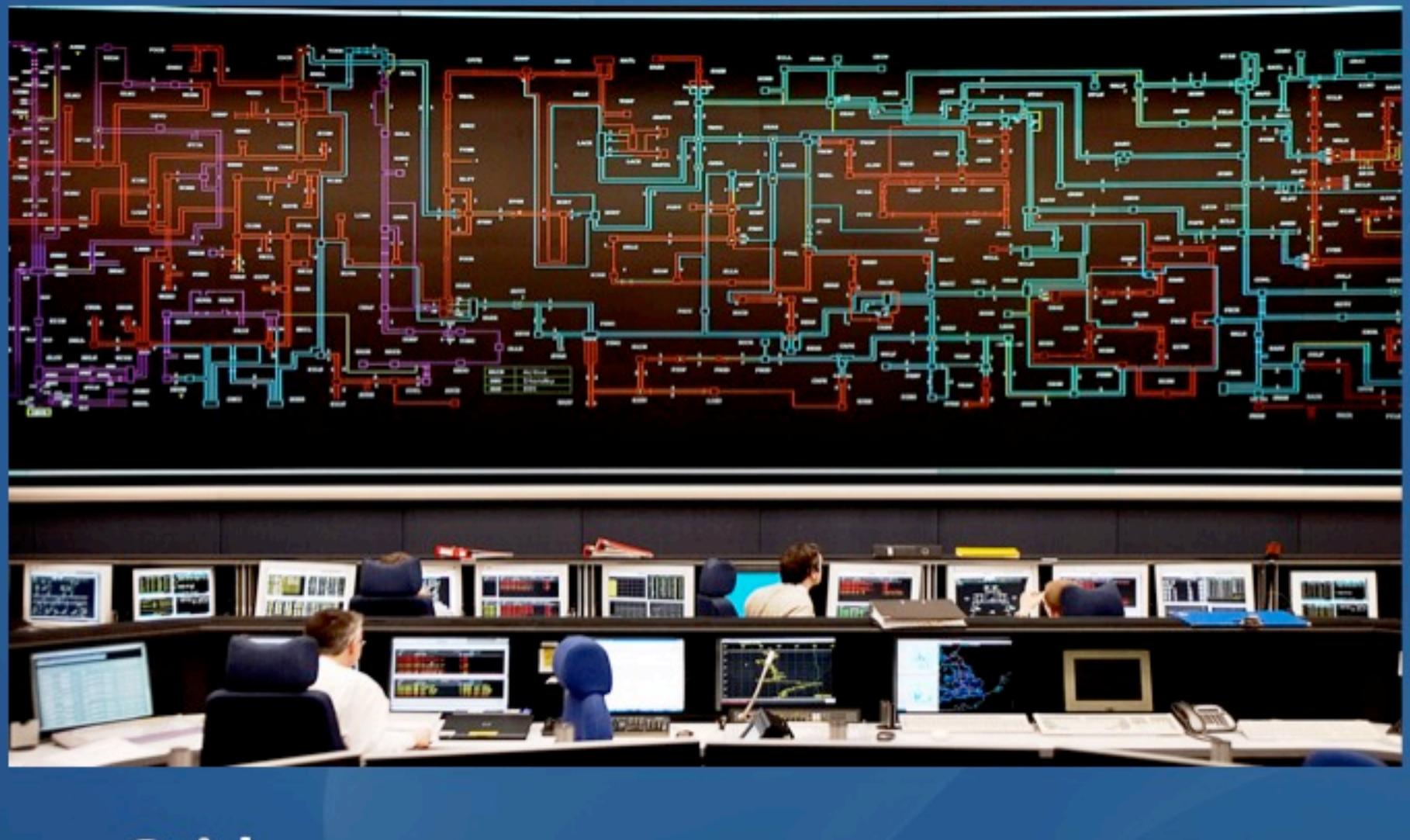
Metrics and Optimum Tests for Sequential Change-Detection, ISyE, GaTech, Dec. 2016

- First-entry: **Model for change-imposing mechanism.**
- Unifies all existing formulations
- Understanding of change-imposing mechanism can explain existing metrics
- May lead to more efficient detectors.

Goal: detect occurrence of τ

τ is a stopping time adapted to the history generated by the (partially) hidden process $\{X_t\}$.

T is a stopping time adapted to the history generated by the observation sequence $\{\xi_t\}$.



Power Grid:

X_t: Energy at major points in the grid.

$\xi_t = X_t + W_t$ noisy measurements.

$\mathcal{A}$: If $X_t \in \mathcal{A}$ then, after short time major blackout.

$\mathcal{A}$ is known

Metrics and Optimum Tests for Sequential Change-Detection, ISyE, GaTech, Dec. 2016



Structure health monitoring:

Dependent

X_t: Vibrations at ~~every~~ point of the structure (state)

$\xi_t = \mathbf{A}X_t + W_t$: Low dimensional noisy measurements

$\mathcal{A}$: If $X_t \in \mathcal{A}$ then cracks (**change** in the structure)

$\mathcal{A}$ known or unknown.

Metrics and Optimum Tests for Sequential Change-Detection, ISyE, GaTech, Dec. 2016







Independent $\{X_t\}$ and $\{\xi_t\}$?:

X_t: Field coordinates of the ball

ξ_t: Noisy vibration measurements

Independent

$\mathcal{A}$: Volume under the goal net.

Performance metrics - known entry set

Delayed Detection

$$\mathcal{J}(T) = \mathsf{E}_1[T - \tau | T > \tau]$$

$$\inf_T \mathcal{J}(T) = \inf_T \mathsf{E}_1[T - \tau | T > \tau]$$

$$\text{subject to : } \mathsf{P}_0(T \leq \tau) \leq \alpha$$

Hard Limited Delay

$$\mathcal{P}(T) = \mathsf{P}_1(T \leq \tau + M | T > \tau)$$

$$\sup_T \mathcal{P}(T) = \sup_T \mathsf{P}_1(T \leq \tau + M | T > \tau)$$

$$\text{subject to : } \mathsf{P}_0(T \leq \tau) \leq \alpha$$

Delayed detection - known entry set

$$\inf_T \mathcal{J}(T) = \inf_T \mathsf{E}_1[T-\tau | T>\tau]$$

$$\text{subject to : } \mathsf{P}_0(T \le \tau) \le \alpha$$

Optimal Stopping Theory

Pair process $\{(X_t\,, \xi_t)\}$ is i.i.d. before and after τ with joint pdfs f_0, f_1.

Moustakides (2016): The optimum test

$$\pi_t = \mathsf{P}_0(X_t \in \mathcal{A} | \xi_t)$$

$$T_o = \inf\{t>0 : S_t \ge \nu\}$$

$$S_t = S_{t-1}\frac{f_1(\xi_t)}{(1-\pi_t)f_0(\xi_t)} + \frac{\pi_t}{1-\pi_t}$$

If additionally $\{X_t\}$ and $\{\xi_t\}$ **independent processes**, then

$$\pi_t = \mathsf{P}_0(X_t \in \mathcal{A}|\xi_t) = \mathsf{P}_0(X_t \in \mathcal{A}) = \pi$$

$$S_t = S_{t-1}\frac{f_1(\xi_t)}{(1-\pi)f_0(\xi_t)} + \frac{\pi}{1-\pi}$$

$$\tilde{S}_t = \frac{1-\pi}{\pi}S_t - 1, \quad \tilde{\nu} = \frac{1-\pi}{\pi}\nu - 1$$

$$T_o = \inf\{t > 0 : \ \tilde{S}_t \geq \tilde{\nu}\}$$

$$\tilde{S}_t = (\tilde{S}_{t-1} + 1)\frac{f_1(\xi_t)}{(1-\pi)f_0(\xi_t)}$$

Shiryayev test (1963)

Unknown entry set

- What if entry set $\mathcal{A}$ is unknown?
- Can detect the first-entry to an unknown set?
  Equivalently: can detect the change-time τ that inflicts a change in the statistical behavior?

Focus on
change of the statistics

Performance metrics - unknown entry set

Delayed Detection

Worst-case analysis

$$\mathcal{J}(T) = \sup_{\tau} \mathsf{E}_1[T - \tau | T > \tau]$$

$$\inf_{T} \mathcal{J}(T) = \inf_{T} \sup_{\tau} \mathsf{E}_1[T - \tau | T > \tau]$$

$$\text{subject to : } \mathsf{E}_0[T] \geq \gamma$$

Hard Limited Delay

$$\mathcal{P}(T) = \inf_{\tau} \mathsf{P}_1(T \leq \tau + M | T > \tau)$$

$$\sup_{T} \mathcal{P}(T) = \sup_{T} \inf_{\tau} \mathsf{P}_1(T \leq \tau + M | T > \tau)$$

$$\text{subject to : } \mathsf{E}_0[T] \geq \gamma$$

Delayed detection - unknown entry set

$$\mathcal{J}(T) = \sup_{\tau} \mathsf{E}_1[T - \tau | T > \tau]$$

We can show (Moustakides 2008):

$\{X_t\}$ and $\{\xi_t\}$ independent processes

$$\mathcal{J}_{\mathsf{P}}(T) = \sup_{t \geq 0} \mathsf{E}_1[T - t | T > t] \quad \text{Pollak (1985)}$$

$\{X_t\}$ and $\{\xi_t\}$ dependent processes

$$\mathcal{J}_{\mathsf{L}}(T) = \sup_{t \geq 0} \sup_{\xi_1, \ldots, \xi_t} \mathsf{E}_1[T - t | T > t, \xi_1, \ldots, \xi_t]$$

Lorden (1971)

$$\mathcal{J}_{\mathsf{P}}(T) = \sup_{t \geq 0} \mathsf{E}_1[T - t | T > t] \quad \text{Pollak (1985)}$$

$$\inf_T J_{\mathsf{P}}(T) \;\; \text{subject to} : \mathsf{E}_0[T] \geq \gamma$$

Discrete time: i.i.d. data before and after the change with pdfs f_0, f_1.

Shiryaev-Roberts-Pollak test

Compute recursively the following statistic:

$$S_t = (S_{t-1} + 1)\frac{f_1(\xi_t)}{f_0(\xi_t)} \quad T_{\mathsf{P}} = \inf\{t > 0 : S_t \geq \nu\}$$

Pollak (1985): If S_0 specially designed, then

$$[J_{\mathsf{P}}(T_{\mathsf{P}}) - \inf_T J_{\mathsf{P}}(T)] \to 0; \quad \text{as } \gamma \to \infty$$

Exact optimality?

~~Yakir (1997)~~ Mei (2007)

Polunchenko-Tartakovsky (2012) counterexample.

Change in the drift of a BM: Polunchenko (2016)

Dependence? Multiple pre- and/or post-change possibilities? Time variation?

$$\mathcal{J}_{\mathsf{L}}(T) = \sup_{t\geq 0} \sup_{\xi_1,\ldots,\xi_t} \mathsf{E}_1[T-t|T>t,\xi_1,\ldots,\xi_t]$$

Lorden (1971)

$$\inf_T J_{\mathsf{L}}(T) \;\; \text{subject to} : \mathsf{E}_0[T] \geq \gamma$$

Discrete time: i.i.d. data before and after the change with pdfs f_0, f_1.

CUSUM test

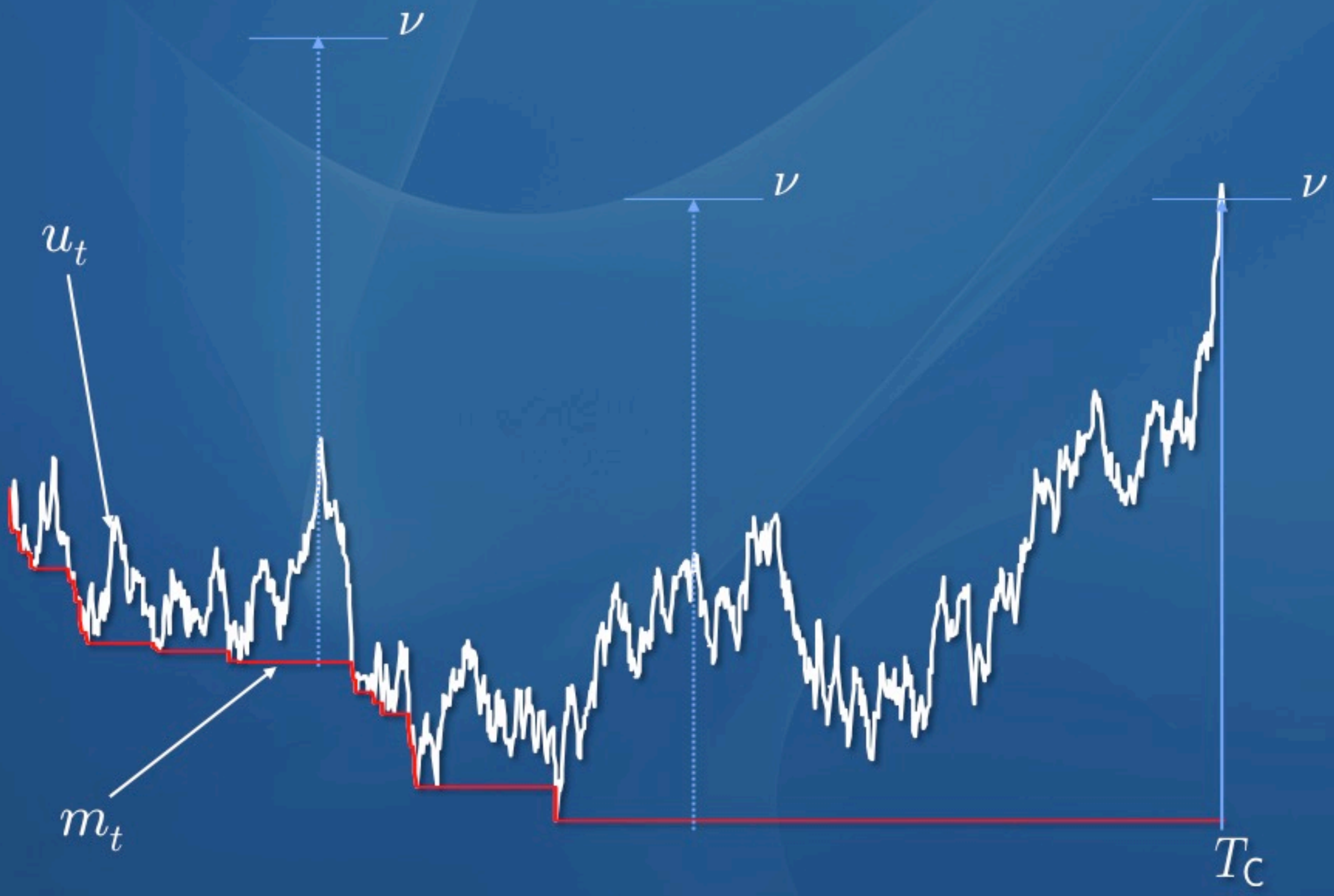
$$S_t = \max\{S_{t-1}, 0\} + \log\frac{f_1(\xi_t)}{f_0(\xi_t)}$$

$$T_{\mathsf{C}} = \inf\{t>0 : S_t \geq \nu\}$$

$$u_t = \sum_{s=1}^{t} \log\frac{f_1(\xi_s)}{f_0(\xi_s)}$$

$$m_t = \min_{0\leq s\leq t} u_s$$

$$S_t = u_t - m_t$$


ν
ν
ν
u_t
m_t
T_C

Discrete time: Lorden (1971) asymptotic optimality.
Moustakides (1986) exact optimality.

Continuous time: Shiryaev (1996), Beibel (1996)
Brownian Motion with constant drift before and after.
Moustakides (2004) strict optimality for Ito processes.

Discrete time: Moustakides-Veeravalli (2016) Non abrupt changes

Dependence? Multiple pre- and/or post-change possibilities?

Hard Limited Delay

$$\mathsf{P}_1(T \leq \tau + M | T > \tau)$$

Only for $M = 1$: $\mathsf{P}_1(T = \tau + 1 | T > \tau)$

Corresponds to immediate detection with the first sample after the change

$\mathcal{P}_{\mathsf{S}}(T) = \mathsf{P}_1(T = \tau + 1 | T > \tau)$ Shiryaev like

$\mathcal{P}_{\mathsf{P}}(T) = \inf_{t \geq 0} \mathsf{P}_1(T = t + 1 | T > t)$ Pollak like

$\mathcal{P}_{\mathsf{L}}(T) = \inf_{t \geq 0} \inf_{\xi_1, \ldots, \xi_t} \mathsf{P}_1(T = t + 1 | T > t, \xi_1, \ldots, \xi_t)$ Lorden like

$$\sup_T \mathcal{P}_{\mathsf{S}}(T) \quad \text{s.t. : } \mathsf{P}_0(T \leq \tau) \leq \alpha$$

$$\sup_T \mathcal{P}_{\mathsf{P(L)}}(T) \quad \text{s.t. : } \mathsf{E}_0[T] \geq \gamma$$

$$T_{\text{Sh}} = \inf\left\{t > 0 : \frac{f_1(\xi_t)}{f_0(\xi_t)} \geq \nu\right\}$$ Shewhart test (1931)

Optimality: Bojdecki (1979): Shiryaev like

Pollak and Krieger (2013): Pollak like

Moustakides (2014): Lorden like

Pollak and Krieger (2013): Multiple post-change possibilities.

Moustakides (2014): Post change time variation

Dependent observations

$$\mathcal{P}_{\mathsf{L}}(T) = \inf_{t\geq 0} \inf_{\xi_1,\ldots,\xi_t} \mathsf{P}_1(T = t+1 | T > t, \xi_1, \ldots, \xi_t)$$

$$\sup_T \mathcal{P}_{\mathsf{L}}(T) \quad \text{subject to : } \mathsf{E}_0[T] \geq \gamma$$

Markovian pre- and post-change observations $\{\xi_t\}$

$$T_{\mathsf{Sh}} = \inf\left\{ t > 0 : c(\xi_{t-1}) \frac{f_1(\xi_t|\xi_{t-1})}{f_0(\xi_t|\xi_{t-1})} \geq \nu(\xi_t) \right\}$$

Moustakides (2015): With properly designed functions $c(\xi)$ and $\nu(\xi)$ we solve the constrained optimization.

Simple solution for conditionally Gaussian pdfs.

$$\mathcal{J}_{\mathsf{L}}(T) = \sup_{t \geq 0} \sup_{\xi_1, \ldots, \xi_t} \mathsf{E}_1[T - t | T > t, \xi_1, \ldots, \xi_t]$$

$$\inf_T J_{\mathsf{L}}(T) \quad \text{subject to} : \mathsf{E}_0[T] \geq \gamma$$

For Markovian pre- and post-change $\{\xi_t\}$: Solution ??

Acknowledgement

NSF: CIF-1513373 through Rutgers University

Collaboration program with UIUC

Thank you for your attention!

Statistics Seminar - George V. Moustakides

Time: Thursday, 1 December 2016 - 12:30pm to 1:30pm

Location: Groseclose 402

Event Details

TITLE: Metrics and optimum tests for sequential change-detection

ABSTRACT:

We consider the problem of sequential detection of a change in the statistical behavior of an observed process. After presenting a panorama of applications from diverse scientific fields that can be formulated as a sequential change detection problem, we introduce the various metrics adopted in the literature for the mathematical setup of the problem of interest. In each case we make a thorough overview of the existing optimum detection strategies, mention applications the formulation is suitable for, and pay particular attention to results that have been recently developed. Finally, we discuss versions of the problems that are still open and have been challenging researchers for many years.

BIO: George V. Moustakides received the diploma in Electrical and Mechanical Engineering from the National Technical University of Athens, Greece, the MSE in Systems Engineering from the Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, USA and the MSc and PhD in Electrical Engineering and Computer Science from Princeton University, Princeton, USA. Since 2007 he is with the department of Electrical and Computer Engineering, University of Patras, Greece while prior to this position he held long-term appointments as junior and senior researcher (directeur de recherche) with INRIA, France and as professor with the University of Thessaly, Greece. During his career Prof. Moustakides also held visiting scholar and/or adjunct professor positions at numerous universities in USA as Princeton University, University of Pennsylvania, Columbia University, University of Maryland, University of Illinois at Urbana-Champaign and recently Rutgers University. He has served as associate editor for Detection and Estimation (2011-2014) and is currently serving as inaugural associate editor for Sequential Methods (2016-2019) for the IEEE Transactions on Information Theory. His research focuses on Sequential Analysis and Statistical Signal Processing.