# A Prior-Free Encode-Decode Change Detection Test to Inspect Datastreams for Concept Drift

Cesare Alippi[12], Li Bu[1], Dongbin Zhao[1]

1. State Key Laboratory of Management and Control for Complex Systems, Institute of Automation Chinese Academy of Sciences
2. Dipartimento di Elettronica e Informazione, Politecnico, di Milano, 20133 Milano, Italy

The International Joint Conference on Neural Networks 2013

# Summary

- The Quest for Change Detection
- Encode-Decode Machine
- Encode-Decode CDT for Detecting Changes
- Experimental Validation
- Conclusions

# The Quest for Change Detection

Fault

Time variance of the environment

Inaccuracy of the change detection mechanism

...

changes in stationarity

**Concept Drift**

**relevant side effects on the application**

classification accuracy drops

# The Quest for Change Detection

**The aim**

to develop a novel and effective change detection test which operates in the residual framework

auto-associative Encode-Decode learning machine

1. hierarchical H-ICI CDT
2. The Lepage CPM

**Encode-Decode CDT**

# Summary

- The Quest for Change Detection
- Encode-Decode Machine
- Encode-Decode CDT for Detecting Changes
- Experimental Validation
- Conclusions

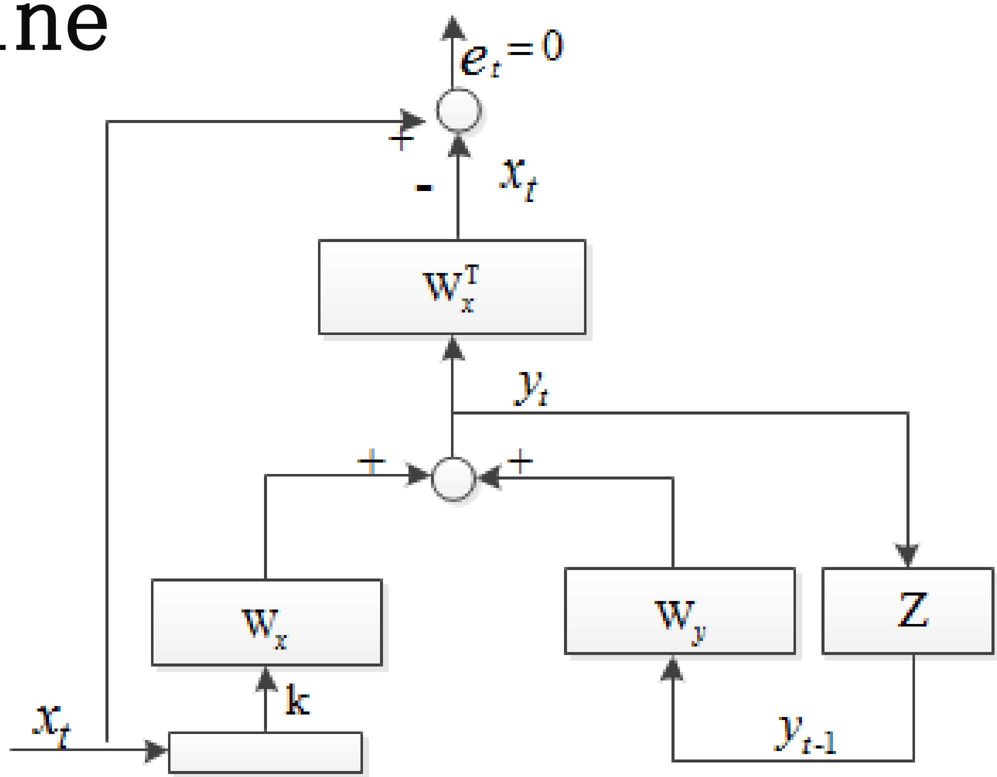
# Encode-Decode Machine

$$y_t = W_x x_t + W_y y_{t-1}$$

satisfy the constraints:

$$x_t = W_x^T y_t$$

$$y_{t-1} = W_y^T y_t$$




Sufficient conditions to design the machine in [1] require:

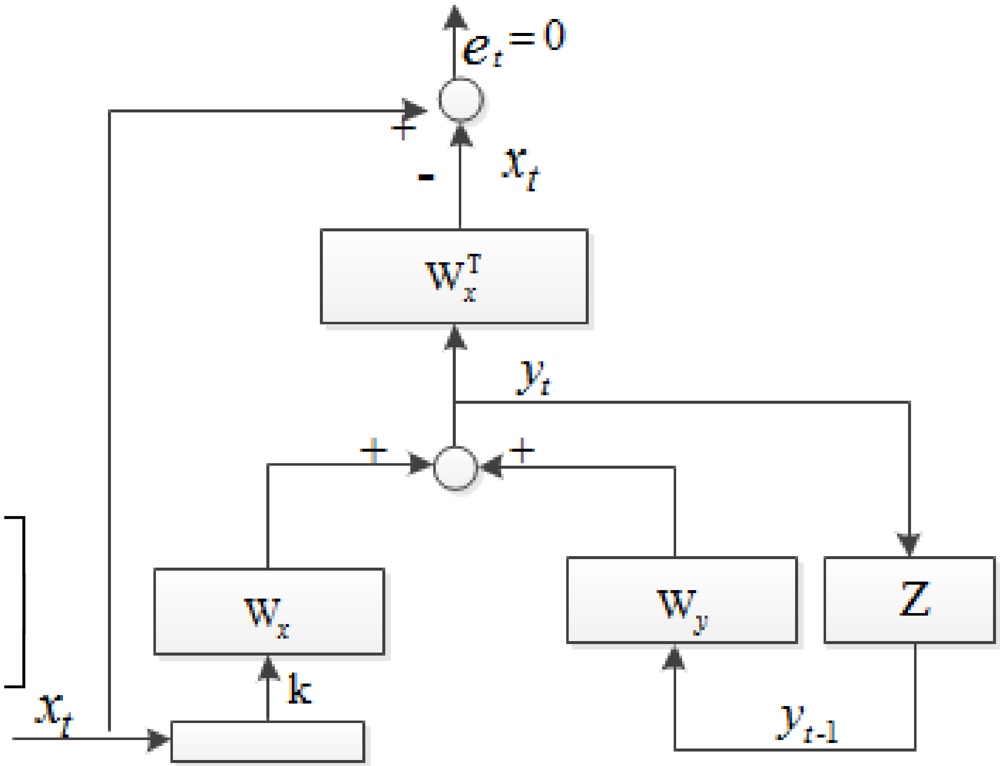
1. the columns of $W_x$ to be orthogonal to those of $W_y$
2. $W_x^T W_x$ and $W_y^T W_y$ must be projector matrices applied to vectors $x_t$ and $y_t$

# Encode-Decode Machine

$y_t$, $W_y$ and $W_x$ are defined as:

$$y_t^T = \left[ x_t^T, x_{t-1}^T, \cdots, \underbrace{0^T, \cdots, 0^T}_{q-t+1} \right]$$

$$W_y = \begin{bmatrix} 0_{k,kq} & 0_{k,k} \\ I_{kq,kq} & 0_{kq,k} \end{bmatrix}; W_x = \begin{bmatrix} I_{k,k} \\ 0_{kq,k} \end{bmatrix}$$




$k$ is the length of vector $x_t$, $q$ is the memory depth of this machine, and define $s = k(q+1)$.

# Encode-Decode Machine

To reduce the complexity, Construct at first matrix $X_{\tau,s}$ for each row vector $y_t$, $t = 1,\dots,\tau$ and compute the SVD

$$X_{\tau,s} = U_{\tau,\tau} \Sigma_{\tau,s} V_{s,s}^T$$

Keep the first $p$ singular values and the associated eigenvetors to compose $\tilde{V}$ .

$$\tilde{V}^T \tilde{V} = I_{p,p}$$

Define $y_t = V^T y_t$, and $V^T W_y y_{t-1} = V^T W_y V V^T y_{t-1}$ when $p \geq rank(\Sigma_{t,s})$

$$\tilde{y}_t = \tilde{V}^T W_x x_t + \tilde{V}^T W_y \tilde{V} \tilde{y}_{t-1}$$

# Encode-Decode Machine

$$\tilde{y}_t = \tilde{V}^T W_x x_t + \tilde{V}^T W_y \tilde{V} \tilde{y}_{t-1}$$

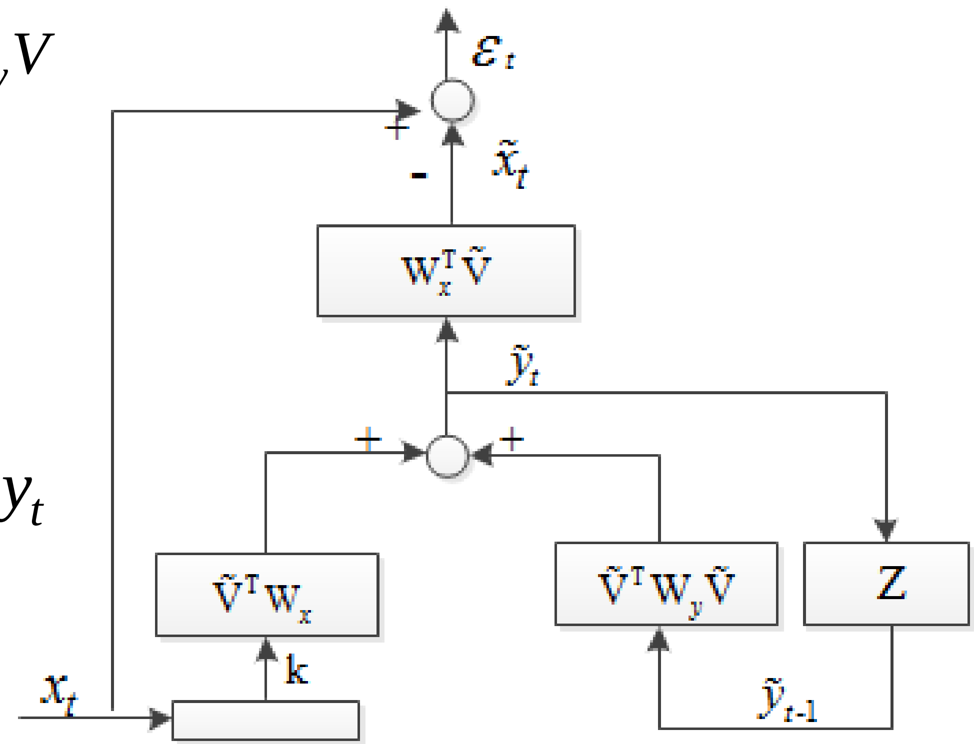
represents an approximate expression when $p < rank(\Sigma_{t,s})$

Define $\mathcal{W}_x = V^T W_x$, $\mathcal{W}_y = V^T W_y V$

$$y_t = \mathcal{W}_x x_t + \mathcal{W}_y y_{t-1}$$

$$\tilde{x}_t = \mathcal{W}_x^T y_t$$

$$\varepsilon_t = x_t - \tilde{x}_t = W_x^T (I_{s,s} - \tilde{V}\tilde{V}^T) y_t$$



# Encode-Decode Machine

Select the $p$ maximizing the whiteness of the reconstruction residual

$$p = \arg\min_{p=1,\dots,s} \sum_{\tau=1}^{T} \left| \sum_{t=1}^{N} \varepsilon_t \varepsilon_{t-\tau} \right|$$

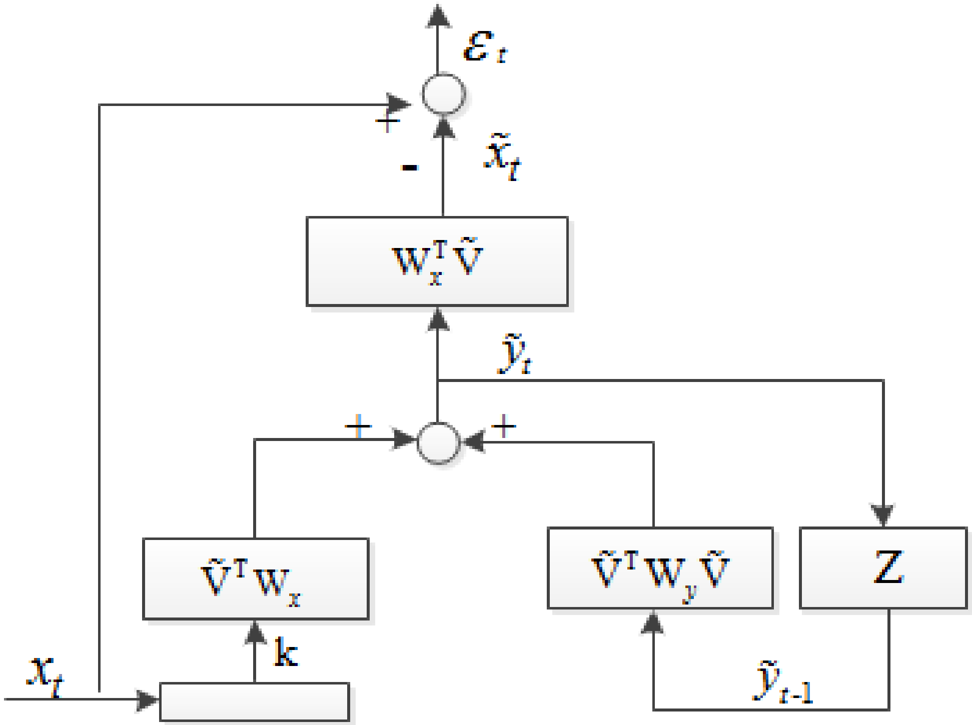



Fig.1 the Encode-Decode SVD-reduced machine

# Summary

- The Quest for Change Detection
- Encode-Decode Machine
- Encode-Decode CDT for Detecting Changes
- Experimental Validation
- Conclusions

# Encode-Decode CDT for Detecting Changes

**A. *Detecting Changes with the Designed Encode-Decode Machine***

$$\varepsilon_t = W_x^T (I_{s,s} - \tilde{V}\tilde{V}^T) y_t = (I_{s,s} - \tilde{V}\tilde{V}^T)_{k,s} y_t$$

the first component of the error vector $\varepsilon_t$

$$e_t = \sum_{i=1}^{s} a_i y_i = a^T y_t$$

$a_i$ is the $i$-th entry of the first row $a$ of the matrix $(I_{s,s} - VV^T)$

# Encode-Decode CDT for Detecting Changes

***(1). Abrupt Perturbation Case***

$$x_\delta(t) = x(t) + \delta,\ t > t_p$$ $\delta(t) = \delta$ is an abrupt perturbation

$$e_{t,\delta} = \sum_{i=1}^{s} a_i y_i + \delta \sum_{i=1}^{s} a_i = e_t + e_\delta$$

***(2). Linear Drift Perturbation Case***

$$x_\delta(t) = x(t) + \delta_d t\ ,\ t > t_p$$ $\delta(t) = \delta_d t$ is the perturbation drift.

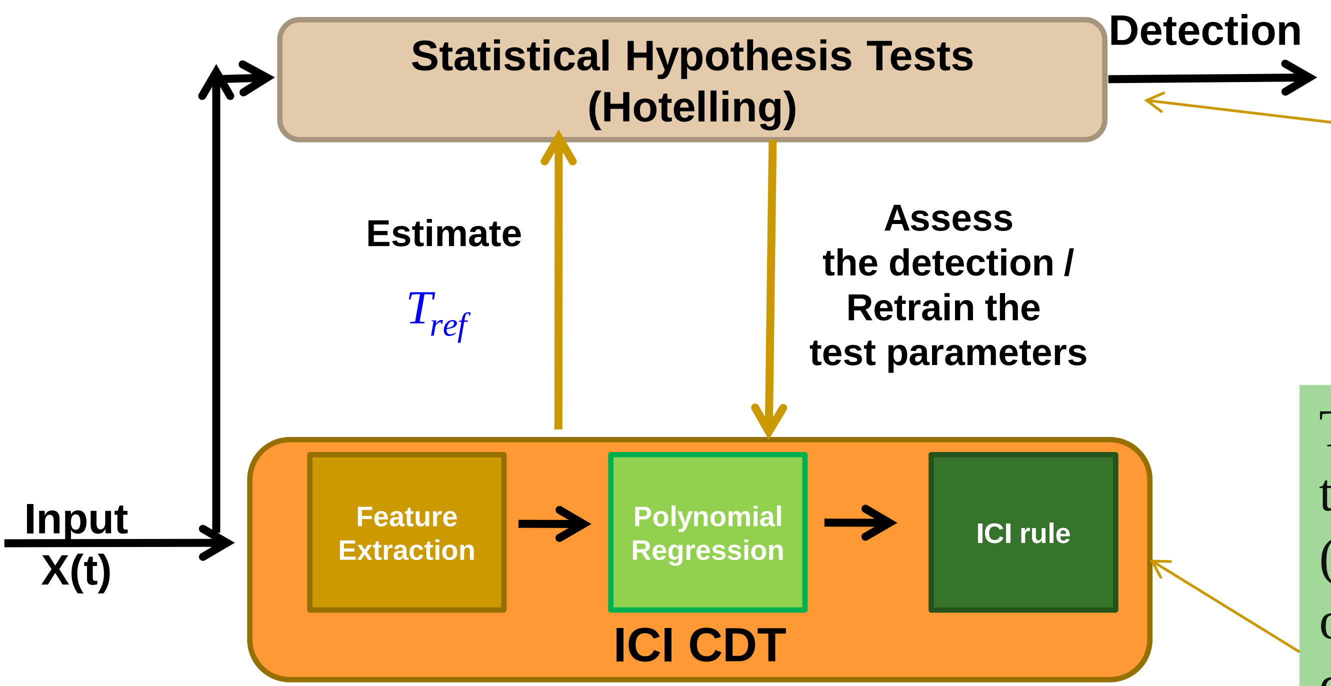
$$e_{t,\delta} = \sum_{i=1}^{s} a_i y_i + \delta \sum_{i=1}^{s} a_i t = e_t + e_\delta$$

# Encode-Decode CDT for Detecting Changes

## B. *H-ICI CDT and Lepage CPM*

### *(1). H-ICI CDT*

**Hierarchical Change Detection**




The second level CDT partitions the datastream in two intervals characterizing the states of the process before and after the supposed change

The first one runs a CDT test providing an alarm (either a real concept drift or a false positive detection) and to the second level CDT.

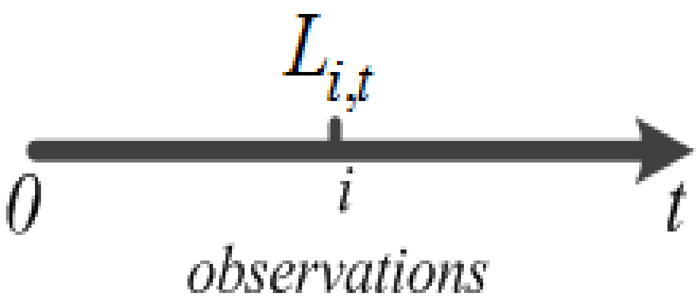
# Encode-Decode CDT for Detecting Changes

**(2). Lepage CPM**

The rank of the *i*-th observation at time *t* is defined as

$$r(x_i) = \sum_{i \neq j}^{t} I(x_i > x_j) \qquad I \text{ is the indicator function}$$

Combine the Mann-Whitney test statistic *U* and the Mood test statistic *M*. $L = U^2 + M^2$ is used.




$$\bar{L} = \max_{0<i\leq t} L_{i,t}$$

a change is detected when $\bar{L} \geq h_t$

( $h_t$ is a given threshold)

# Encode-Decode CDT for Detecting Changes

## C. *Designing the Encode-Decode CDT*

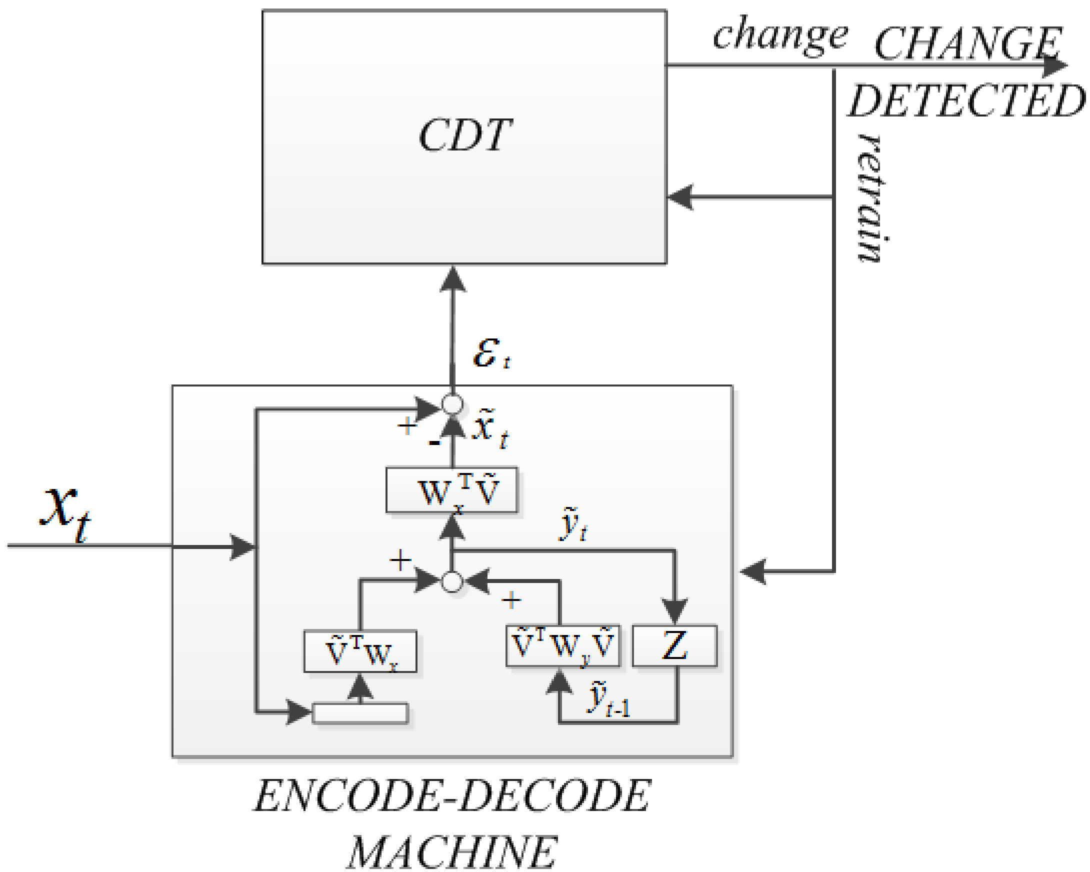



Fig.2 Encode-Decode machine

# Summary

- The Quest for Change Detection
- Encode-Decode Machine
- Encode-Decode CDT for Detecting Changes
- Experimental Validation
- Conclusions

# Experimental Validation

The reference application comes from the data acquired from the hybrid wired-wireless monitoring system deployed on the Alps to monitor a potential rock collapse.

Data, here temperature, are sampled every 5 minutes and sent to a remote control station for subsequent processing.

---

Training set: 3000 samples $k = 1, q = 3$

Validation set: 1000 samples $p = 2$

Testing set: change is injected randomly around sample 1000

Parameters of CDT: gamma.H-ICI=0.5; ARL0.CPM=40000

# Experimental Validation

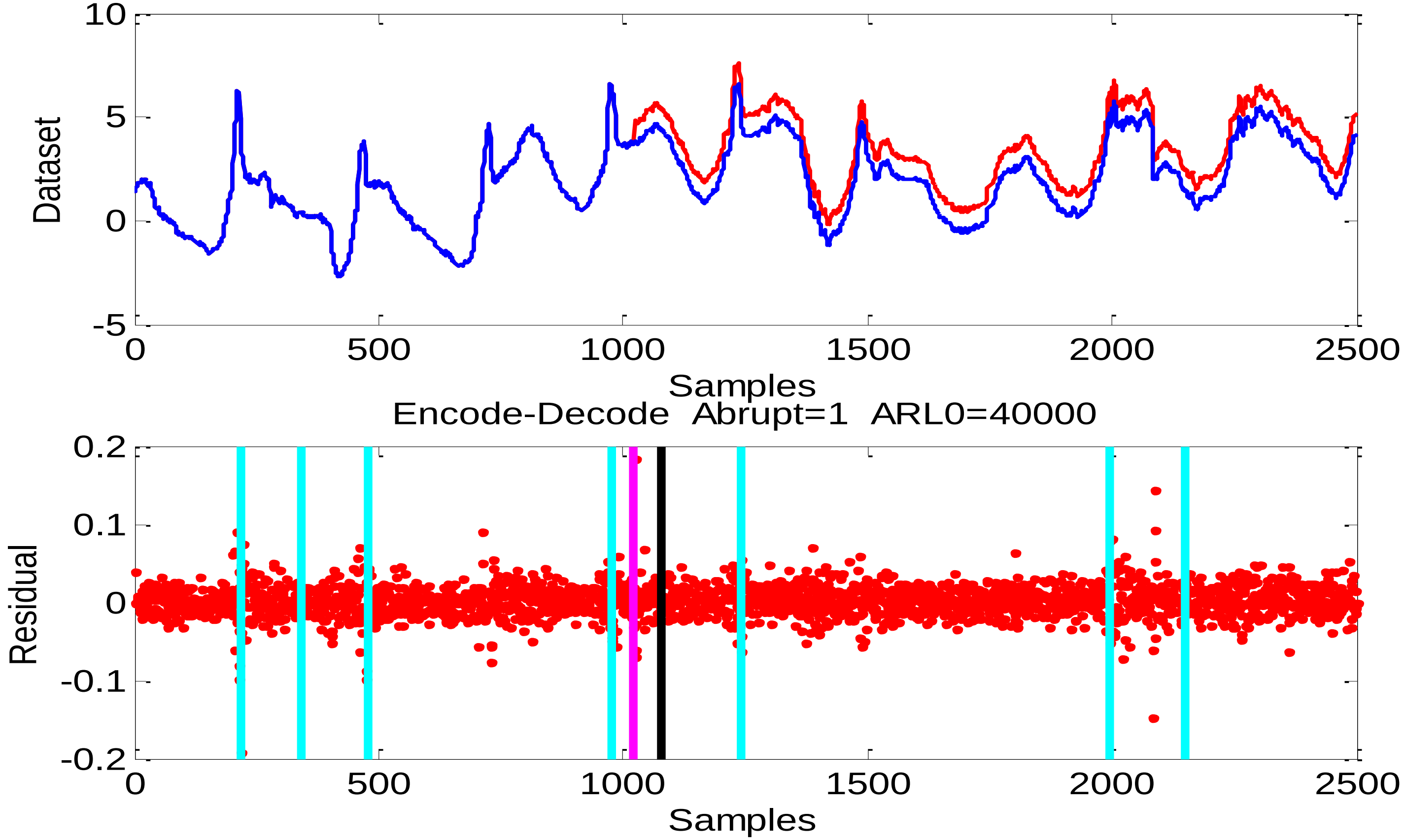



Fig. 3. An abrupt change occurring around point 1000 affects the sensor measurements. (In the upper plot, the blue line represents the initial data, and the red line shows the data with the added change. In the lower plot, the pink vertical line represents the time of the change occurrence, the black one represents the detection of the H-ICI CDT and the cyan one that of the CPM-LP test)

# Experimental Validation

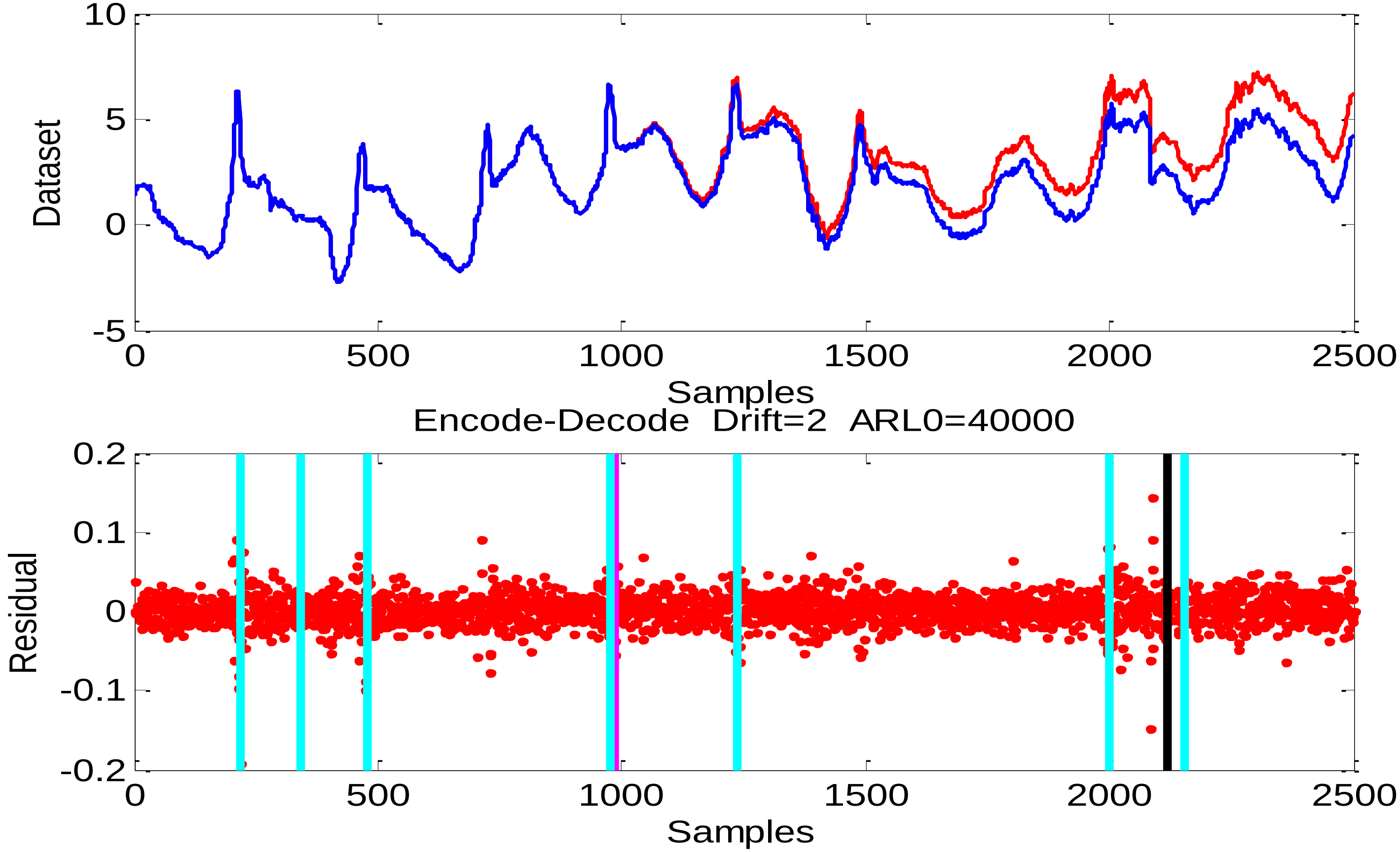



Fig. 4. A linear drift affects the sensor measurements. (In the upper plot, the blue line represents the initial data, and the red line shows the data with the added change. In the lower plot, the pink vertical line represents the start time of change occurrence, the black one is the detection of the H-ICI CDT and the cyan line that of the CPM-LP)

# Experimental Validation

TABLE I The Performance Comparison Of The Encode-Decode Machine And AR With H-ICI And CPM-LP CDT

| Change | *d* | Encode-Decode machine + H-ICI | | | AR + H-ICI | | | Encode-Decode machine + CPM-LP | | | AR + CPM-LP | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | FP | FN | L | FP | FN | L | FP | FN | L | FP | FN | L |
| Abrupt | 0.1 | 0% | 100% | - (-) | 0% | 100% | - (-) | 100% | 0% | 214.8 (65.4) | 100% | 0% | 208.3 (57) |
| | 0.4 | 0% | 59% | 76.7 (50.7) | 0% | 100% | - (-) | 100% | 0% | 213.3 (67) | 100% | 0% | 621 (404.8) |
| | 0.5 | 0% | 25% | 85.2 (54.5) | 0% | 100% | - (-) | 100% | 0% | 201.9 (77.1) | 100% | 0% | 648.4 (404.8) |
| | 1 | 0% | 6% | 84.3 (35.8) | 0% | 7% | 1077.1 (161.7) | 100% | 0% | 199.7 (85.3) | 100% | 0% | 875.3 (307.8) |
| Drift | 0.5 | 0% | 100% | - (-) | 0% | 100% | - (-) | 100% | 0% | 190.5 (84.8) | 100% | 0% | 207.2 (54) |
| | 1 | 0% | 100% | - (-) | 0% | 100% | - (-) | 100% | 0% | 200.3 (79.9) | 100% | 0% | 199.1 (68.4) |
| | 1.5 | 0% | 0% | 1341.5 (20.1) | 0% | 100% | - (-) | 100% | 0% | 197 (83.7) | 100% | 0% | 208.6 (57.5) |
| | 2 | 0% | 0% | 1119.7 (17.5) | 0% | 100% | - (-) | 100% | 0% | 208.2 (72) | 100% | 0% | 209.4 (54.1) |
| | 20 | 0% | 0% | 218.4 (115.7) | 0% | 0% | 1057.4 (23.2) | 100% | 0% | 190.4 (84.8) | 100% | 0% | 207.1 (54) |

# Summary

- The Quest for Change Detection
- Encode-Decode Machine
- Encode-Decode CDT for Detecting Changes
- Experimental Validation
- Conclusions

# Conclusions

- The paper presents a new mechanism for generating a residual sequence to be inspected for potential concept drift.
- The method is based on an auto-associative recurrent machine whose computational complexity is controlled with a SVD decomposition.
- The outcome is a novel change detection test we name Encode-Decode CDT.
- The method, applied to a real application shows the effectiveness.

*Thanks!*