# Statistical Process Monitoring of Nonlinear Profiles Using Wavelets

June 4, 2007

QPRC

## Problem

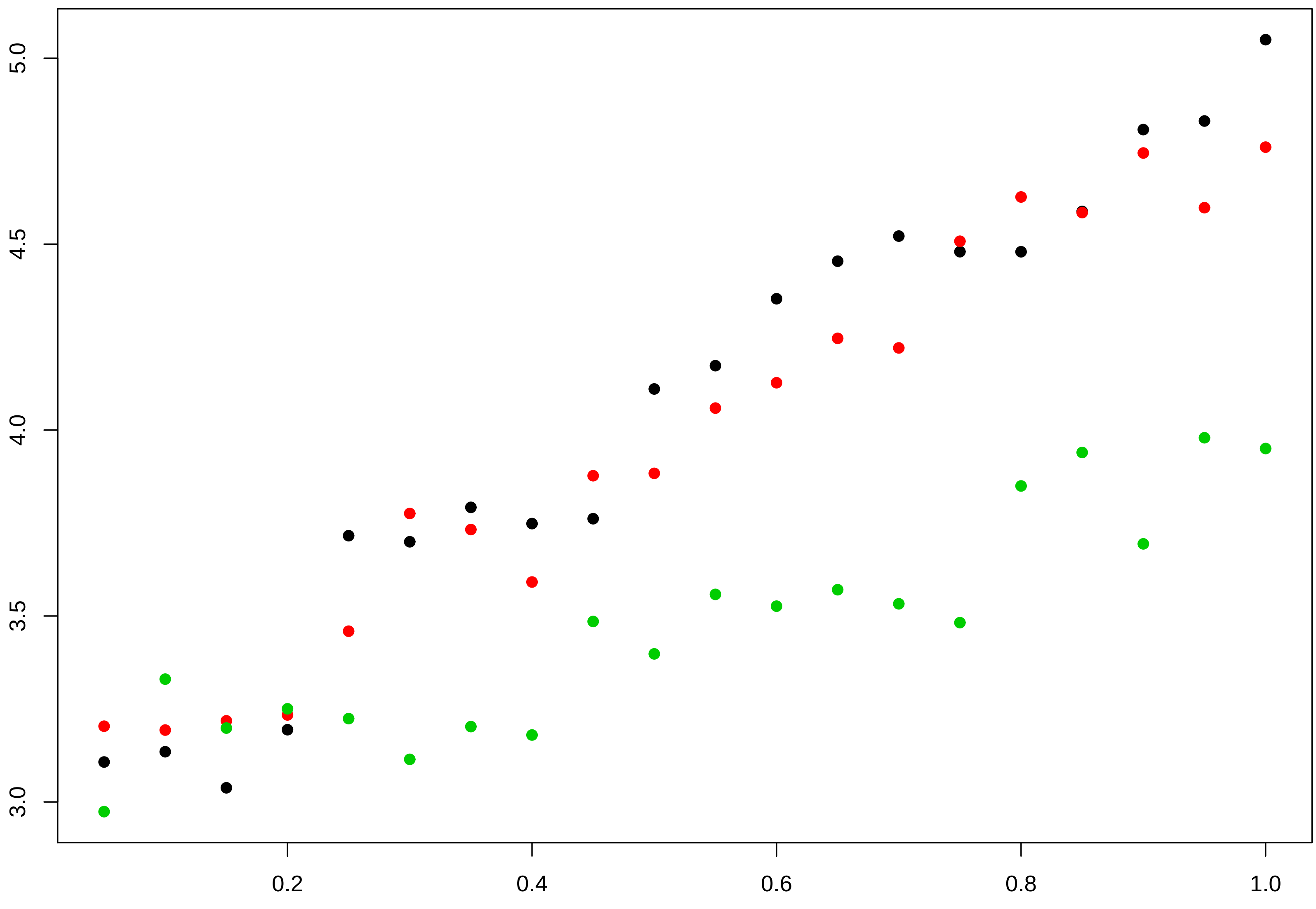
- Modern industrial processes generate complex data
- Profiles: data pairs $(x, y)$ that can be described as $y = f(x)$
- Examples
  - Calibration curves in chemical processing
  - Oxide thickness across wafers in semiconductors
  - Radar signals of military targets.
- Examine sequences of such data sets

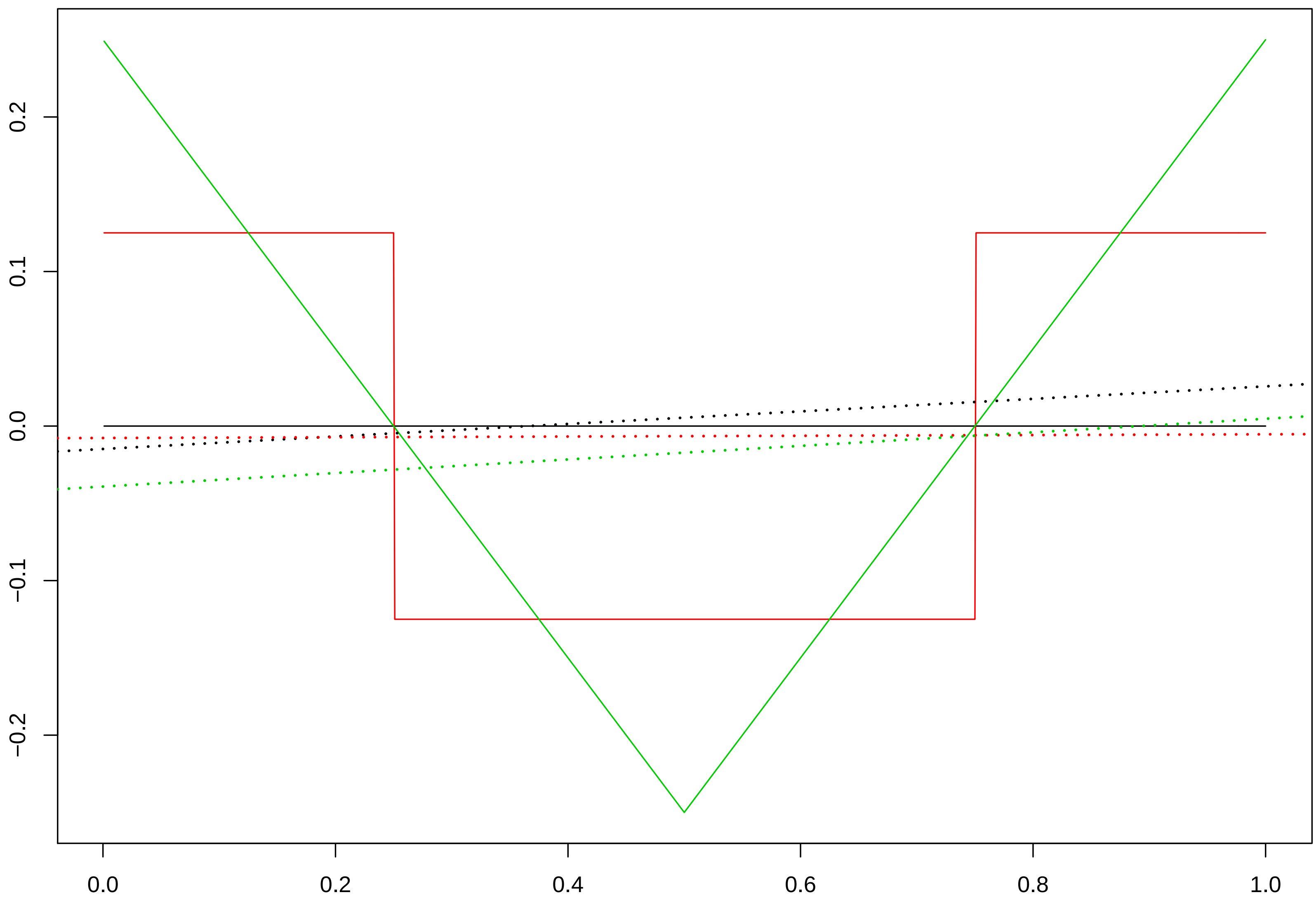
## Problem

- Want to know if a profile is different from some desired, in-control state
- When did change occur?
- What is the nature of change?
- Want a method that applies to very general profile functions $f$

## Problem

- Much work on linear profiles (Woodall 2004, Mahmoud et al. 2007)
- Estimate changes in parameters (slope, intercept, . . .)
- Determine when parameters change

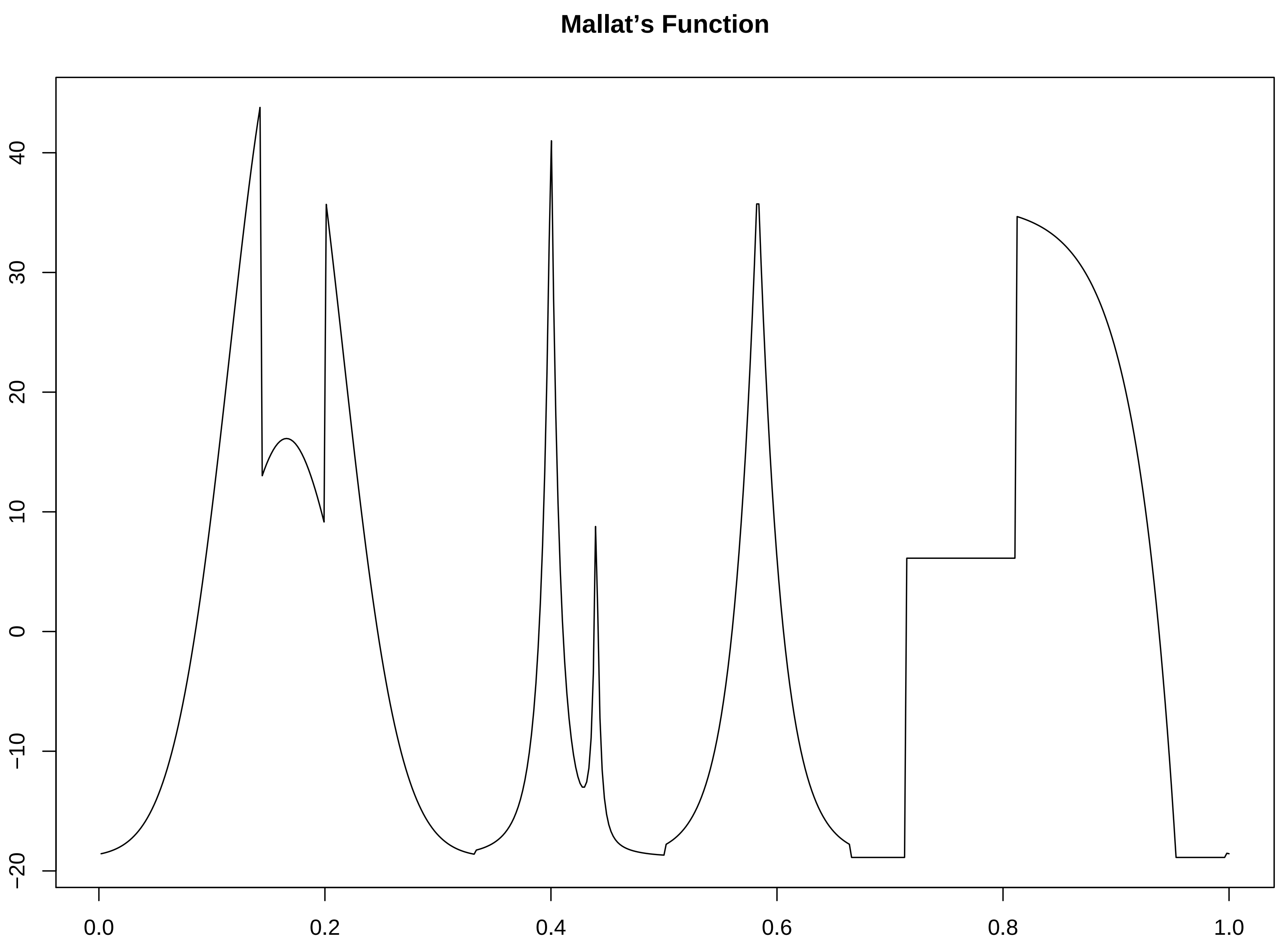

5.0
4.5
4.0
3.5
3.0
0.2
0.4
0.6
0.8
1.0

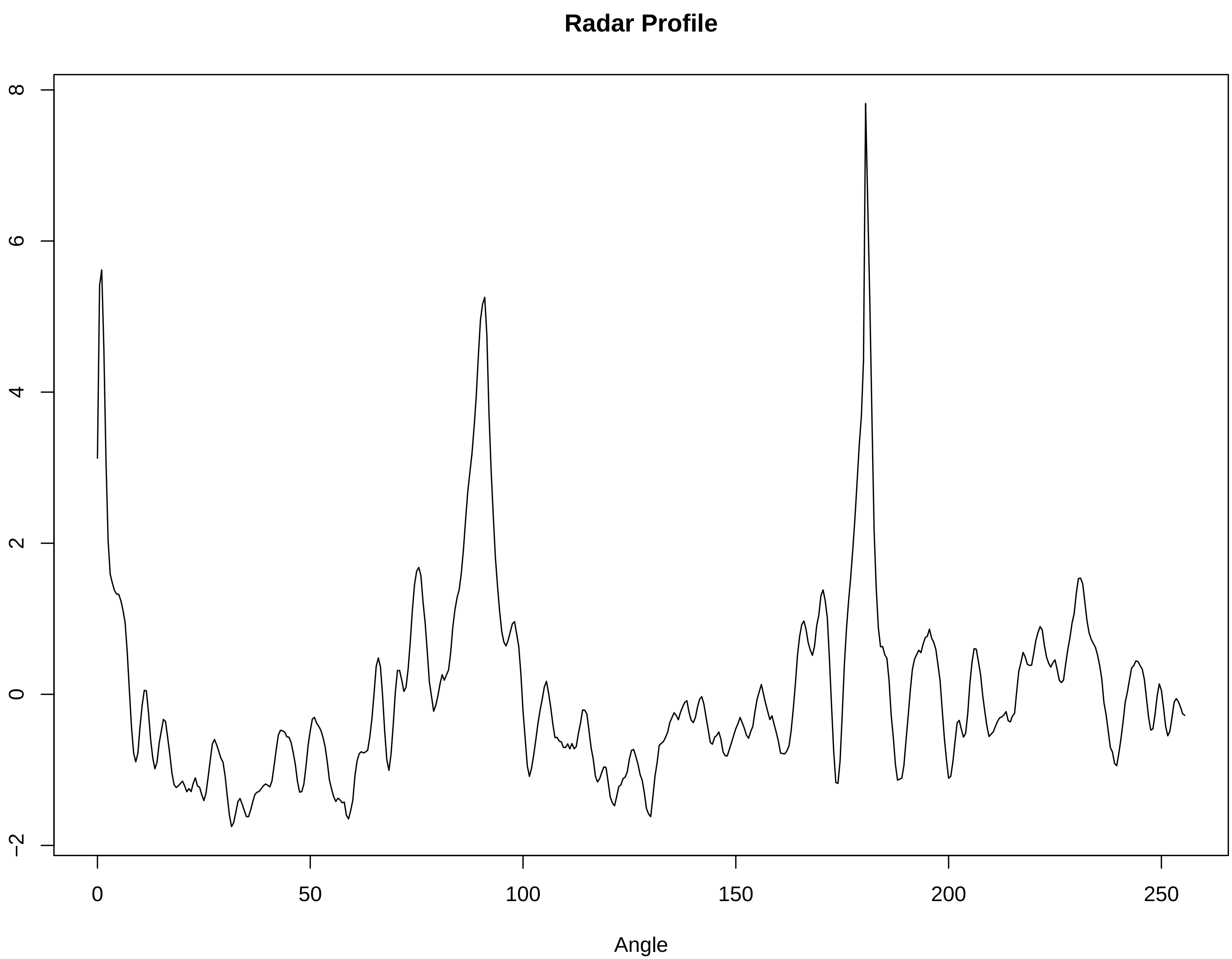


0.2
0.1
0.0
−0.1
−0.2
0.0
0.2
0.4
0.6
0.8
1.0

## Problem

- What if profiles are not linear?
- Or more generally, not parametric?


Mallat's Function
40
30
20
10
0
−10
−20
0.0
0.2
0.4
0.6
0.8
1.0



Radar Profile
8
6
4
2
0
−2
0
50
100
150
200
250
Angle

## Problem

- $f^0$ is known, in-control profile, $f^t$ is observed profile
- Suppose $f^t$ is a very general function:

$$\|f^t - f^0\|_2^2 = \int (f^t - f^0)^2 < \infty$$

- No constraint on form of the profile
- Very weak constraint on the difference of profiles

Formulation of Problem

- Observed profiles

$$y^t = f^t(x) + \epsilon$$

- $\epsilon \sim$ normal $(0, \sigma^2)$, independent
- Hypotheses

$$H_0 : \ \|f^t - f^0\|_2^2 = 0, \qquad t = 1, 2, \ldots, T$$

$$H_a : \ \|f^t - f^0\|_2^2 > 0, \qquad t = \tau + 1, \tau + 2, \ldots, T$$

- In particular

$$f^0 = f^1 = \cdots = f^\tau \neq f^{\tau+1} = \cdots = f^T$$

- Find $\tau$

## Formulation of Problem

- These $L_2$ differences can be written in terms of wavelets coefficients

$$\|f^t - f^0\|_2^2 = \|\theta^t - \theta^0\|_2^2$$

- Moved from function domain to wavelet domain
- Why wavelets?
  - Well-suited for nonparametric estimation
  - Don't need to know much about the form of the functions being estimated
  - Optimally small estimation errors (Donoho & Johnstone 1994)
  - Fast computation time
  - Good at local and global estimation simultaneously

## Formulation of Problem

- Rewrite hypotheses in terms of wavelets

$$H_0: \ \|\theta^t - \theta^0\|_2^2 = 0 \text{ for } t = 1, 2, \ldots, T$$

$$H_a: \ \|\theta^t - \theta^0\|_2^2 > 0 \text{ for } t = \tau + 1, \tau + 2, \ldots, T$$

- Use observed (noisy) data and Discrete Wavelet Transform (DWT) to estimate differences

$$\|f^t - f^0\|_2^2 = \|\theta^t - \theta^0\|_2^2 \approx \|\tilde{\theta}^t - \tilde{\theta}^0\|_2^2$$

## Formulation of Problem

- Set

$$W_t = \frac{n}{\sigma^2}\|\tilde{\theta}^t - \tilde{\theta}^0\|_2^2$$

- Then, for each $t$, $W_t \sim \chi^2_{n,\gamma}$ where

$$\gamma = \frac{n}{\sigma^2}\sum_j (\theta_j^t - \theta_j^0)^2 = \frac{n}{\sigma^2}\|\theta^t - \theta^0\|_2^2$$

  is the non-centrality parameter

- Equivalent hypotheses:

  $H_0 : \; \gamma = 0$ for $t = 1, 2, \ldots, T$

  $H_a : \; \gamma > 0$ for $t = \tau + 1, \tau + 2, \ldots, T$

- Form a likelihood using $W_t$, $t = 1, 2, \ldots, T$
- Under the null hypothesis,

$$L_0 = \prod_{t=1}^{T} f(w_t) = \prod_{t=1}^{T} \frac{w_t^{n/2-1} e^{-w_t/2}}{2^{n/2}\Gamma(n/2)}.$$

- Under the alternative,

$$\begin{aligned} L_a &= \prod_{t=1}^{\tau} \frac{w_t^{n/2-1} e^{-w_t/2}}{2^{n/2}\Gamma(n/2)} \\ &\quad \cdot \prod_{t=\tau+1}^{T} \left\{ \frac{w_t^{n/2-1} e^{-w_t/2}}{2^{n/2}} \sum_{k=0}^{\infty} \frac{e^{-\gamma/2}(\gamma/4)^k}{k!} \cdot \frac{w_t^k}{\Gamma(n/2+k)} \right\} \end{aligned}$$

## Formulation of Problem

- The likelihood ratio can be expressed as

$$\frac{L_a}{L_0} = \prod_{t=\tau+1}^{T} \left\{ \sum_{k=0}^{\infty} \frac{e^{-\gamma/2}(\gamma/4)^k w_t^k}{k!} \cdot \frac{\Gamma(n/2)}{\Gamma(n/2+k)} \right\}$$

- Simplified, the log of the likelihood ratio is

$$\log\left(\frac{L_a}{L_0}\right) \approx \frac{\gamma}{2} \sum_{t=\tau+1}^{T} \left( \frac{w_t}{E(W_t|H_0)} - 1 \right)$$

- Need to estimate $\gamma$

# Test

- Estimate $\gamma$ with DWT and thresholding
- Wavelets are sparse: they concentrate the information in a function into relatively few coefficients
- So, most coefficients can be treated as 0
- Thresholding sets to 0 (or shrinks toward 0) select coefficients $\theta$
- Gives accurate estimation
- Removes noise

- Let $\hat{\gamma}(\tau)$ be thresholded wavelet estimate of $\gamma$

$$\hat{\gamma}(\tau) = \frac{1}{T-\tau}\sum_{t=\tau+1}^{T} \|\hat{\theta}_d^t\|_2^2 - \frac{1}{\tau}\sum_{t=1}^{\tau} \|\hat{\theta}_d^t\|_2^2$$

- Depends on unknown $\tau$

## Test

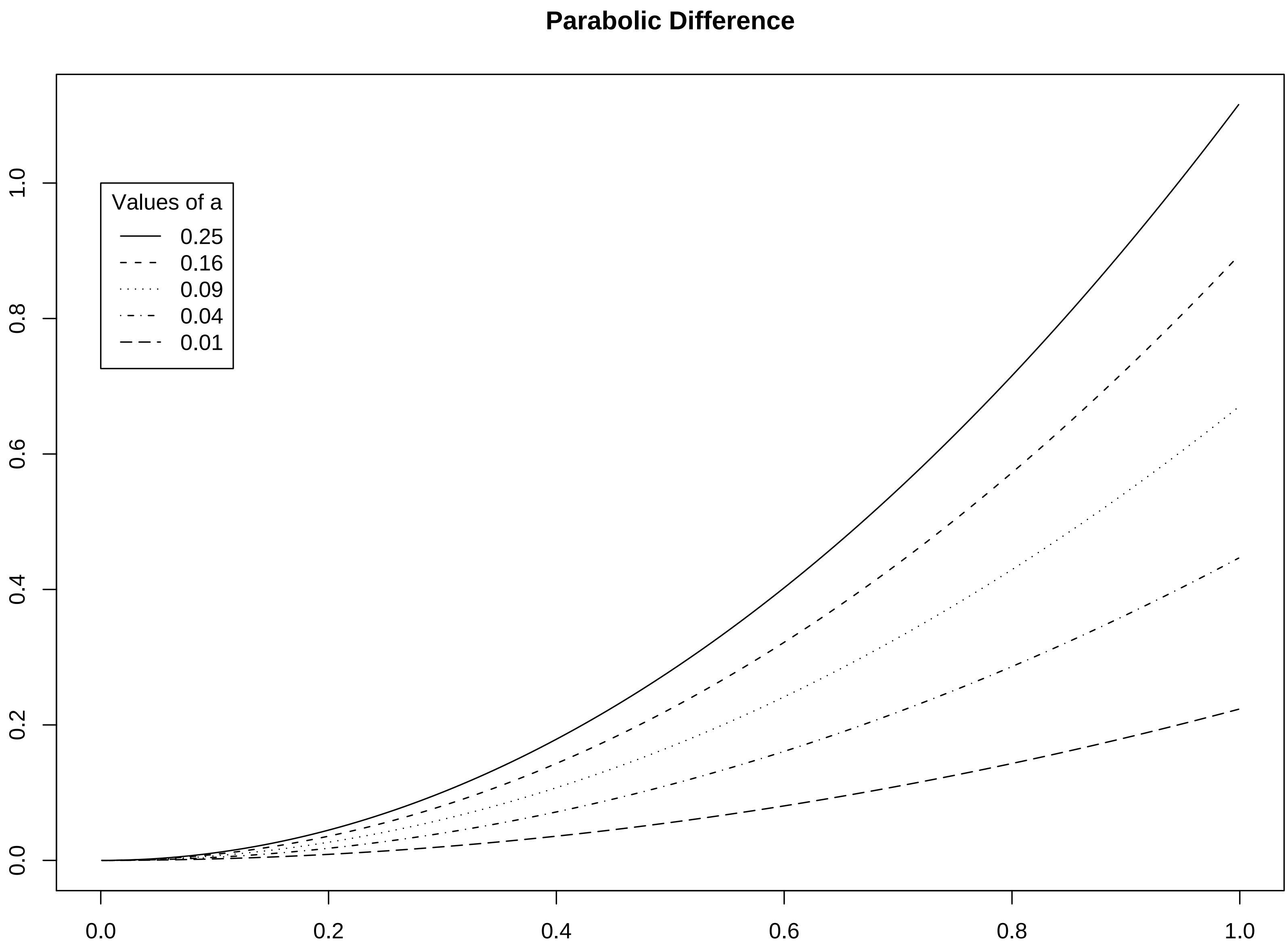
- Reject $H_0$ when likelihood ratio is large
- When is $\log(L_a/L_0)$ largest? $\Rightarrow$ When $\tau$ is correctly specified
- So, maximize $\log(L_a/L_0)$ over $\tau$
- Provides estimate of $\tau \Rightarrow$ estimate of $\gamma \Rightarrow$ estimate of LR

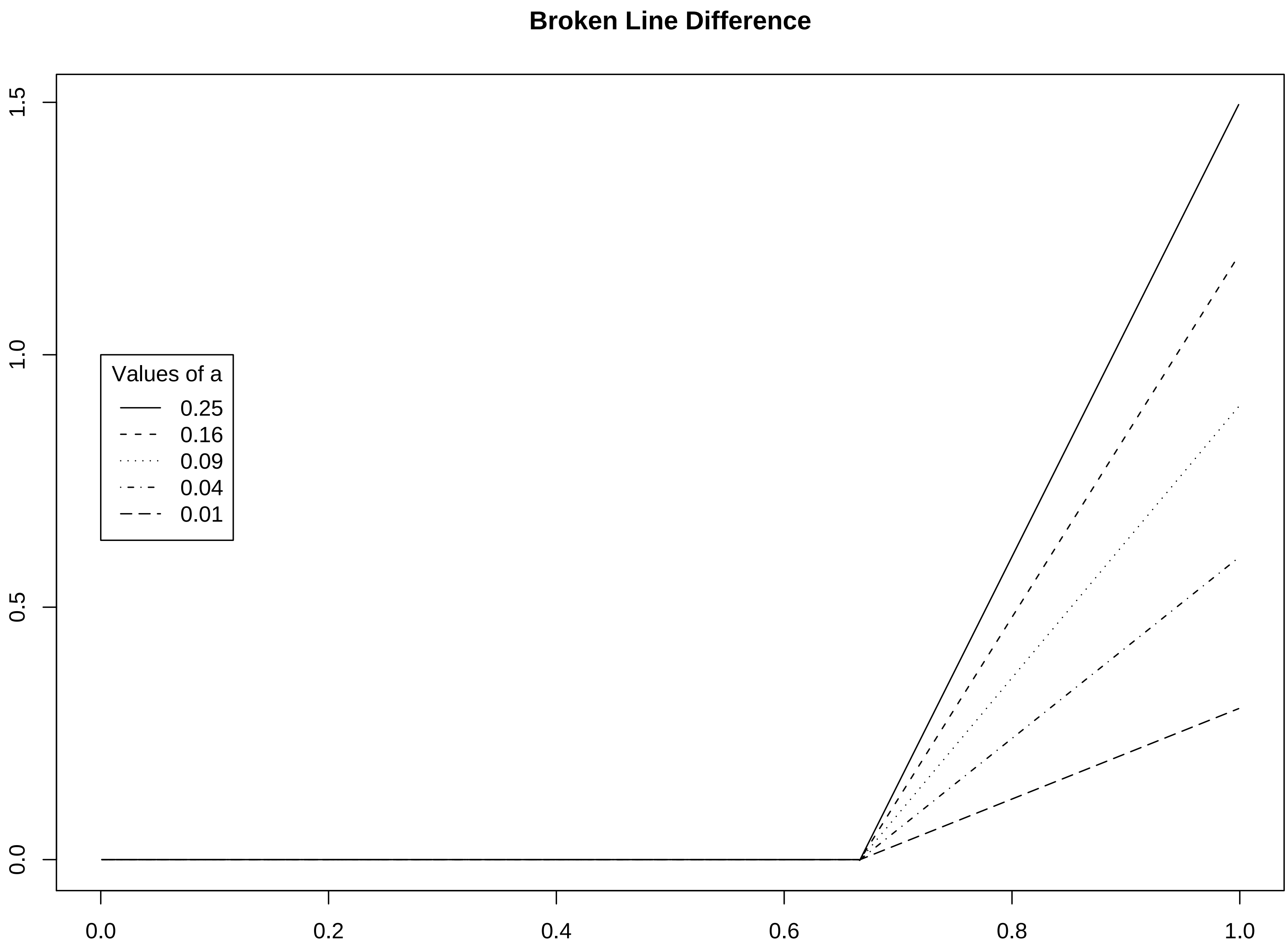
## Test

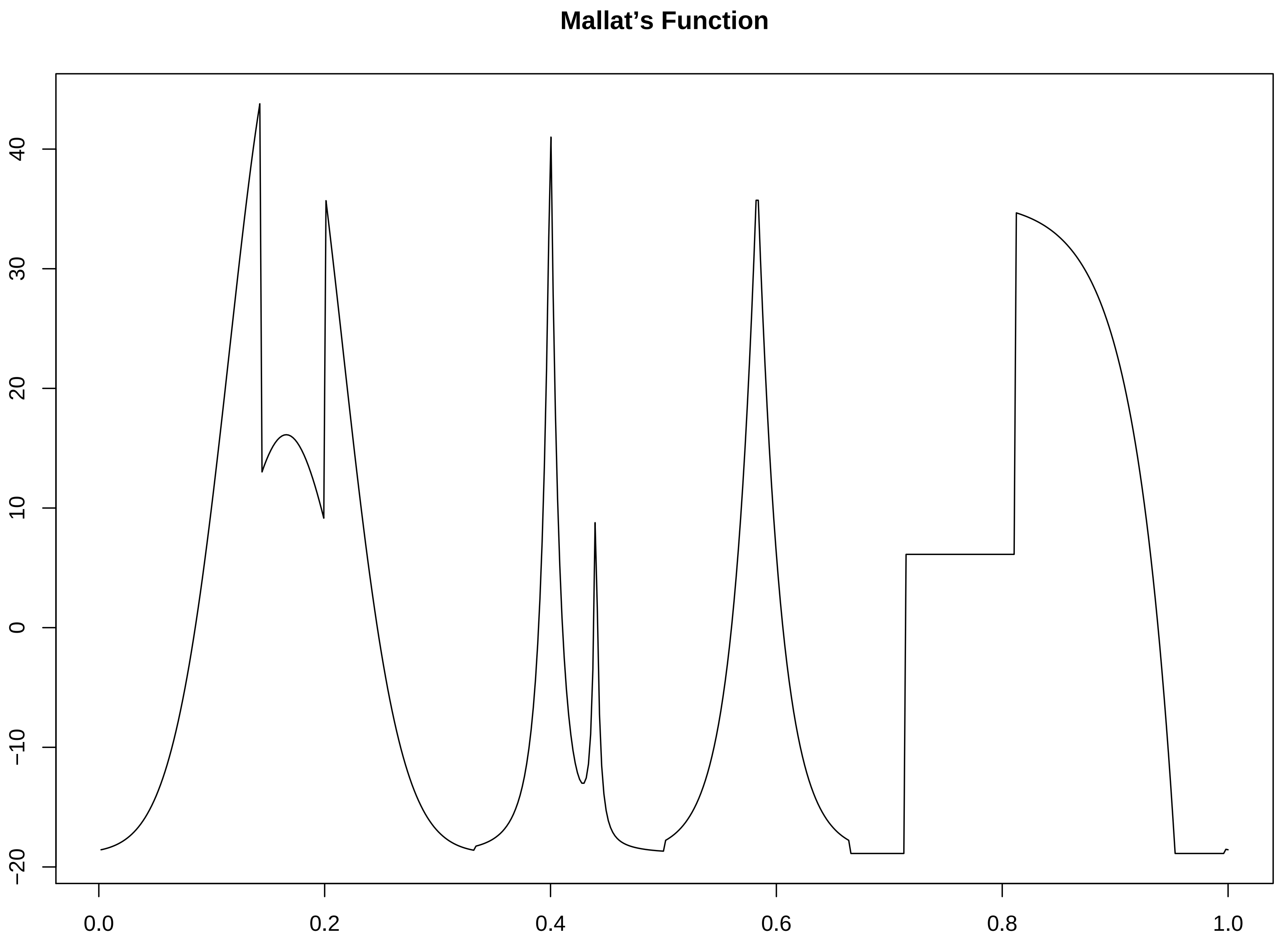
- Estimate of $\tau$ uses prior information: all profiles up to the current profile are observed
- Use this to determine if LR is "too large"
- "Too large" is found via simulation ($ARL_0 = 200$)
- If LR large, then profiles after $\tau$ are out-of-control
- Otherwise, profiles are still in-control at $T$

# Simulation

- This proposed method works well, even for very small differences ($L_2$ difference)
- Simulated several types of difference functions
  - Parabolic
  - Horizontal shift
  - Broken Line
  - Isolated linear shifts
  - Others
- Used erratic profile


Parabolic Difference
Values of a
0.25
0.16
0.09
0.04
0.01
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0


Broken Line Difference
Values of a
0.25
0.16
0.09
0.04
0.01
0.0
0.5
1.0
1.5
0.0
0.2
0.4
0.6
0.8
1.0


Mallat's Function
40
30
20
10
0
−10
−20
0.0
0.2
0.4
0.6
0.8
1.0

## Simulation

- Compare to $M^1$ (Fan 1996), $M^2$ (Jin & Shi 2001), $M^3$ (Jeong et al. 2006)
- Compare via $ARL_0$
- Three wavelet based estimators
- These three do not provide $\tau$ or size of divergence
- Do not use prior information, either

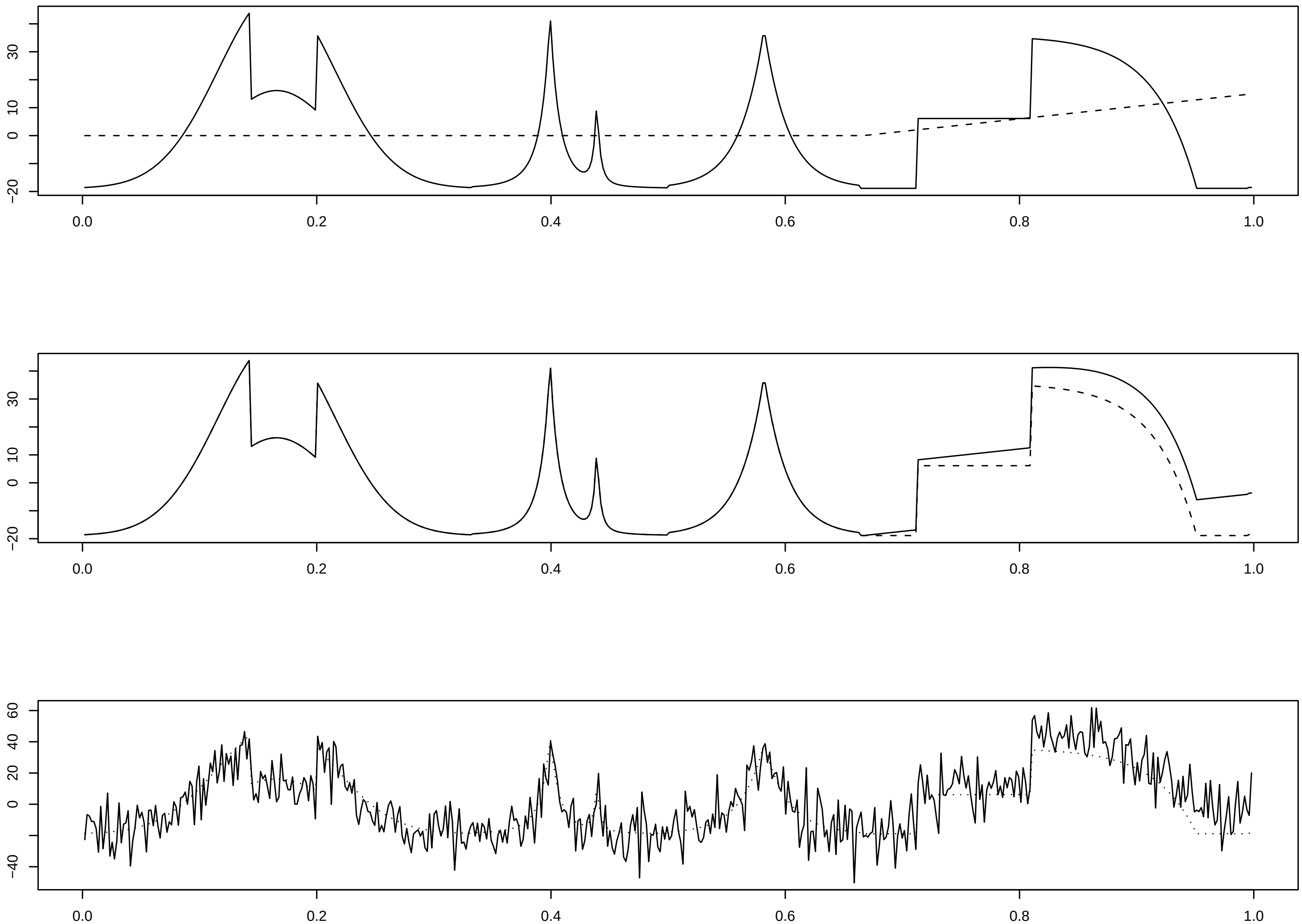
| Parabolic | 0.01 | 0.04 | 0.09 | 0.16 | 0.25 |
|---|---|---|---|---|---|
| $M^1$ | 124.87 | 36.84 | 8.02 | 2.18 | 1.14 |
| $M^2$ | 88.07 | 14.64 | 2.69 | 1.16 | 1.01 |
| $M^3$ | 65.68 | 21.89 | 6.96 | 2.12 | 1.14 |
| $M^*$ | 74.18 | 5.29 | 1.39 | 1.02 | 1.00 |
| Broken Line | 0.01 | 0.04 | 0.09 | 0.16 | 0.25 |
| $M^1$ | 126.80 | 37.28 | 8.38 | 2.30 | 1.18 |
| $M^2$ | 91.84 | 14.26 | 2.70 | 1.14 | 1.00 |
| $M^3$ | 69.36 | 19.52 | 6.38 | 2.12 | 1.17 |
| $M^*$ | 85.57 | 8.83 | 1.69 | 1.03 | 1.00 |

| Parabolic | 0.01 | 0.04 | 0.09 | 0.16 | 0.25 |
|---|---|---|---|---|---|
| $ARL_{10}$ | 76.07 | 5.30 | 1.32 | 1.01 | 1.00 |
| $\hat{\tau}$ | 50.13 | 11.84 | 9.69 | 9.86 | 9.99 |
| $\hat{a}$ | 0.03 | 0.05 | 0.08 | 0.15 | 0.25 |
| Broken Line | 0.01 | 0.04 | 0.09 | 0.16 | 0.25 |
| $ARL_{10}$ | 86.62 | 7.94 | 1.67 | 1.03 | 1.00 |
| $\hat{\tau}$ | 55.33 | 12.79 | 9.99 | 9.86 | 9.99 |
| $\hat{a}$ | 0.03 | 0.04 | 0.08 | 0.15 | 0.25 |

$$\tau = 10$$

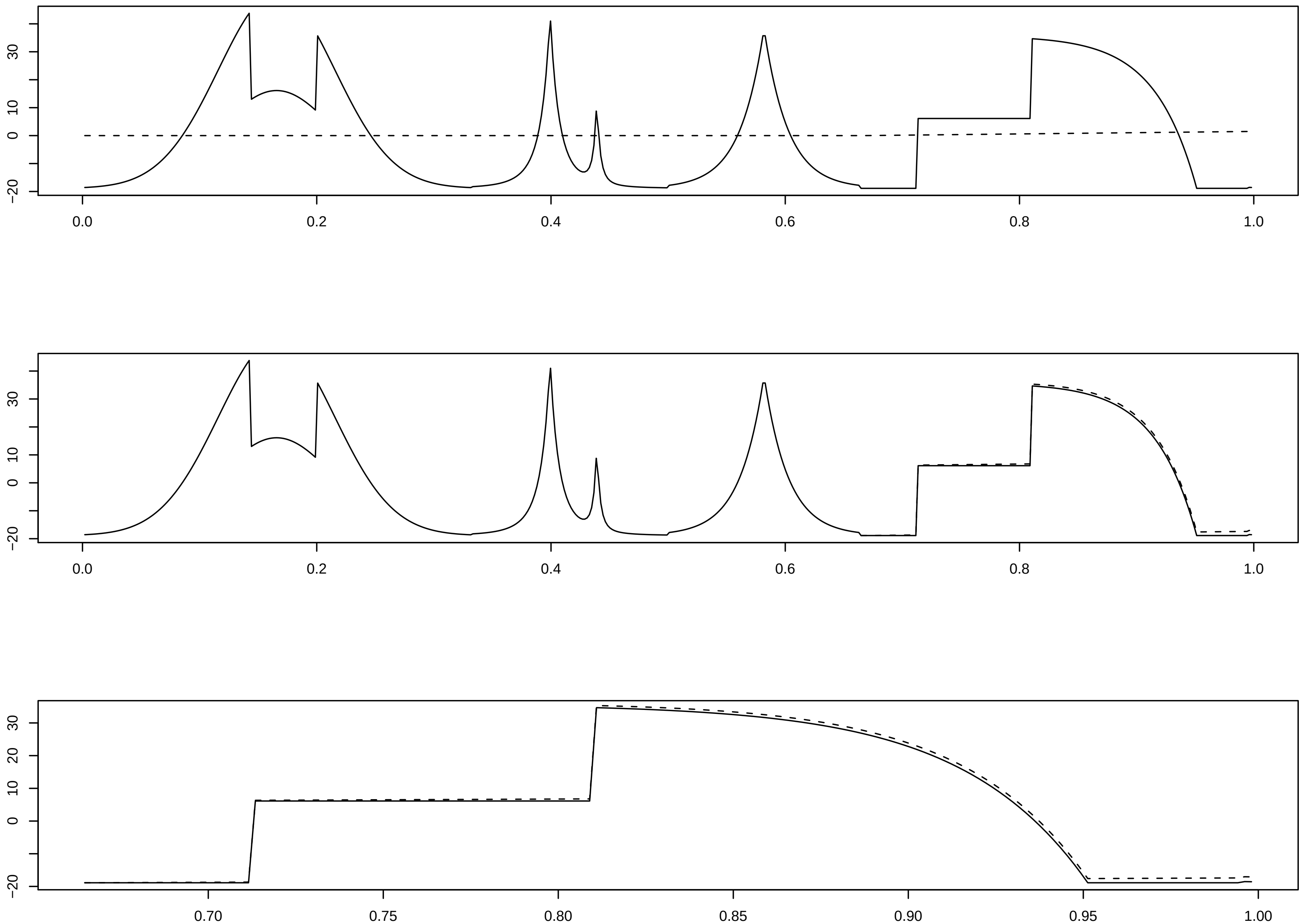
## Simulation

- What is the scale of these differences we are detecting?
- On next graph, exaggerate the "Broken Line" difference by 100 ($a = 25$)


-20
0
10
30
0.0
0.2
0.4
0.6
0.8
1.0
-20
0
10
30
0.0
0.2
0.4
0.6
0.8
1.0
-40
0
20
40
60
0.0
0.2
0.4
0.6
0.8
1.0

# Simulation

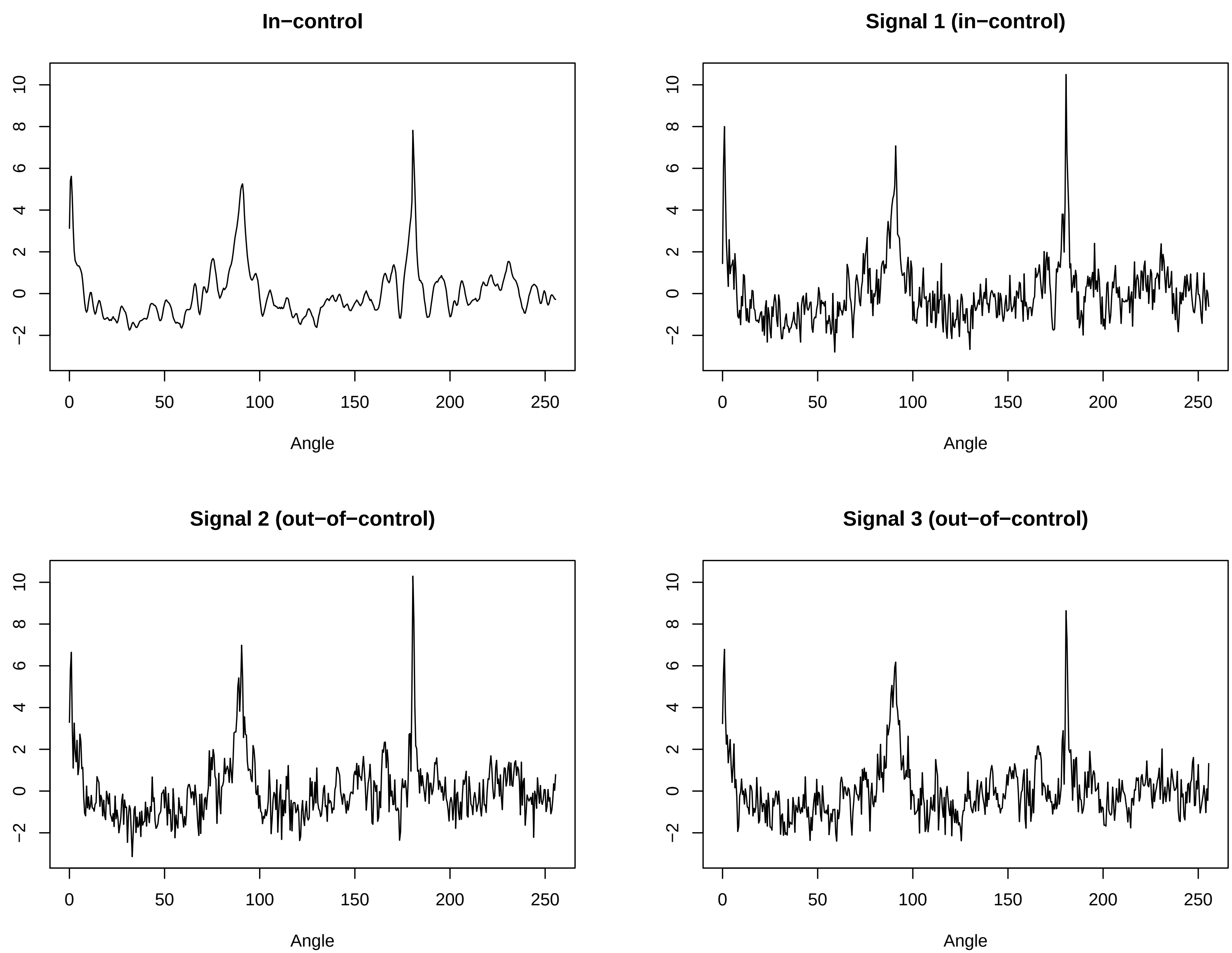
- What was actually looked at was $a \leq 0.25$
- Next graph, $a = 0.25$, the largest difference considered



−20
0
10
30
0.0
0.2
0.4
0.6
0.8
1.0
−20
0
10
30
0.0
0.2
0.4
0.6
0.8
1.0
−20
0
10
20
30
0.70
0.75
0.80
0.85
0.90
0.95
1.00

## Radar Profiles

- Proposed method now applied to profiles of military radar signatures
- Changes from a known profile of the target could indicate
  - A vehicle had moved
  - Some new ground activity was taking place
- Proposed method correctly identified out-of-control profile for any ordering of profiles


In−control
Signal 1 (in−control)
Signal 2 (out−of−control)
Signal 3 (out−of−control)
Angle
0
50
100
150
200
250
−2
0
2
4
6
8
10

## Summary

- Examine functions in wavelet domain
- Form likelihood ratio
- Use wavelet thresholding to estimate parameters in the LR
- Proposed methods specifies when to reject $H_0$
- Tells when out-of-control profile occurred ($\hat{\tau}$)
- Estimates amount of divergence from in-control
- Makes use of prior information
- Joint work with J. Simpson & J. Pignatiello (FSU, IE Dept)

# References


Donoho, D. and Johnstone, I. (1994). Ideal spatial adaptation via wavelet shrinkage. *Biometrika* **81** 425–455.

Fan, J. (1996). Test of significance based on wavelet thresholding and Neyman's truncation. *J. Am. Stat. Assoc.* **91** 674–688.

Jeong, M., Lu, J. and Wang, N. (2006). Wavelet-based SPC procedure for complicated functional data. *International Journal of Production Research* **44** 729–744.

Jin, J. and Shi, J. (2001). Automatic feature extraction of waveform signals for in-process diagnostic performance improvement. *J. Intell. Manuf.* **12** 140–145.

Mahmoud, M., Parker, P., Woodall, W. and Hawkins, D. (2007). A change point method for linear profile data. *Qual. Reliab. Eng. Int.* **23** 247–268.

Woodall, W., Spitzner, D., Montgomery, D. and Gupta, S. (2004). Using control charts to monitor process and product quality profiles. *Journal of Quality Technology* **36** 309–320.