# Materials 38 - Bias in the neighborhood of zero forecast errors

Laura Gáti

July 21, 2020

## Overview

# 1 Simulated "true" data

### 3 potential causes to lack of identification in the zero neighborhood

1. The distribution of estimates is skewed → take $median(\hat{\alpha})$ instead of the mean.

2. The truth is based on a simulation that doesn't favor the zero neighborhood → do 100 simulations from the "true" parameters and take the mean moments of those.

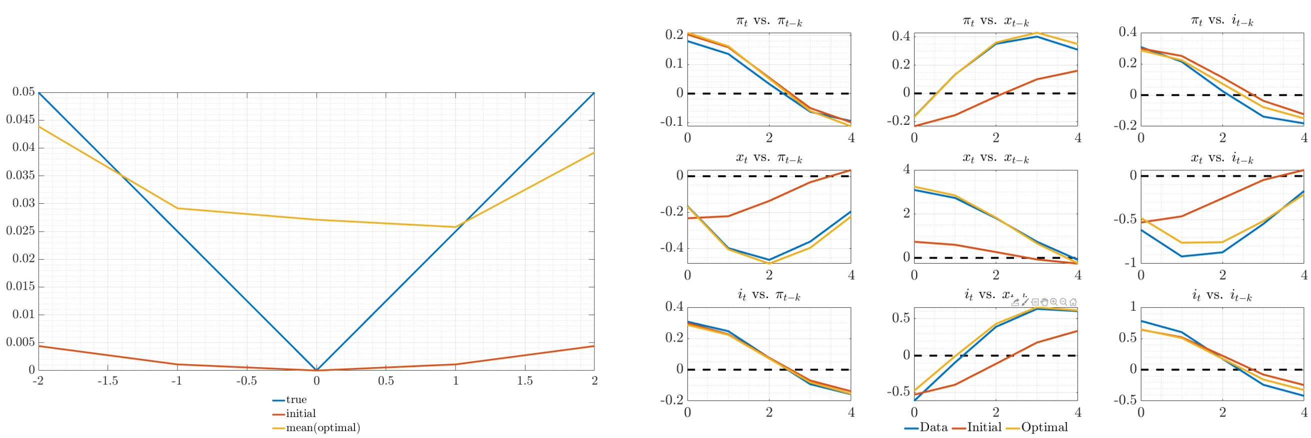
3. The gain doesn't matter if the forecast error is 0, or very close to it → introduce a distinction between the forecast error that's used to choose the gain and the one used to update the coefficients of the learning rule.

+1 Taking mean moments across $N$ histories is more natural than performing the estimation $N$ times.

+2 Introduce expectation series (SPF)

### Reference for comparison: Fig 1. of Materials 37

**Figure 1:** Reference figure: Mean estimates for $N = 100$, imposing convexity with weight 100K, truth with $nfe = 5, fe \in (-2, 2)$

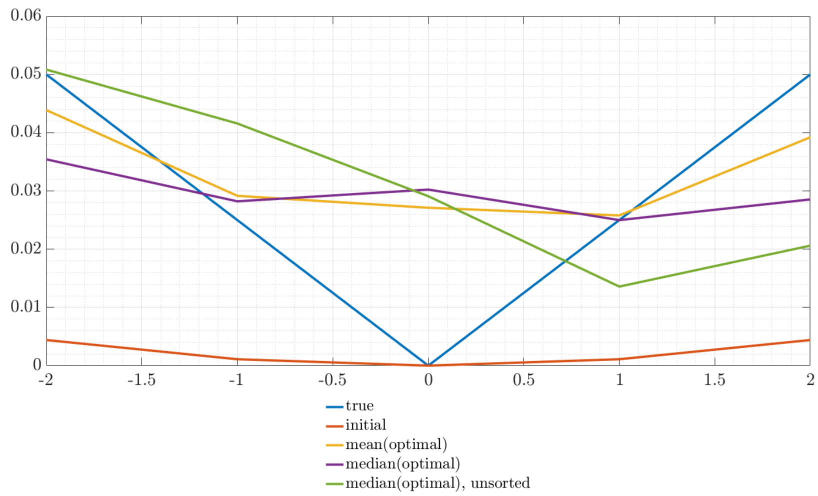



(a) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$

(b) Autocovariogram

## Point #1: skewness → take median instead of mean

**Figure 2:** Mean estimates for $N = 100$, imposing convexity with weight 10K, truth with $nfe = 5, fe \in (-2, 2)$

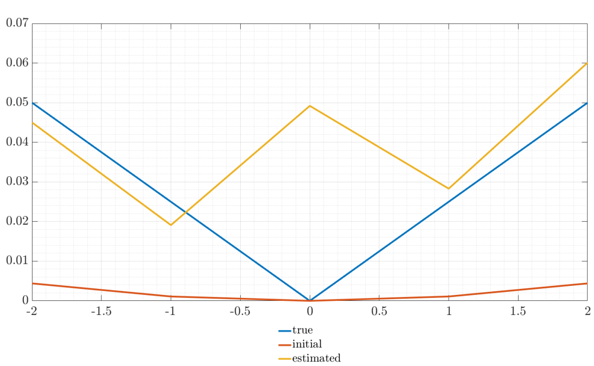



(a) $\alpha^{true}, \alpha_0, mean(\hat{\alpha}), median(\hat{\alpha})$, unsorted median

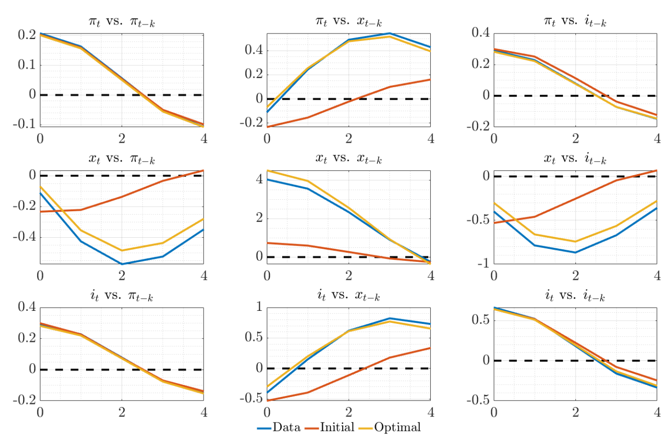
I understand what's happening! Half the estimates are L's, the other half are "inverted L's", which is why taking a mean or a classical, sorted median has the tendency to produce these nonmonotonic zigzags.

## Point #2: do 100 truths

**Figure 3:** Estimates for $N = 100$, truth is a mean of 100 simulations, imposing convexity with weight 100K, truth with $nfe = 5, fe \in (-2, 2)$




(a) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$




(b) Autocovariogram

That didn't help, did it now?

### Point #3: change timing of forecast errors

$$k_t^{-1} = \mathbf{g}(fe_{t|t-1}) \tag{1}$$

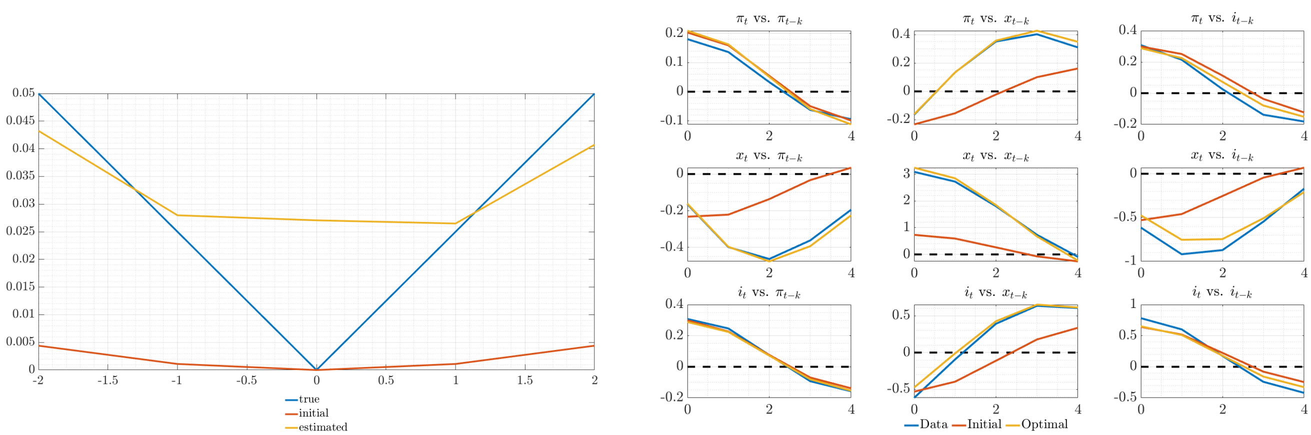
$$\bar{\pi}_t = \bar{\pi}_{t-1} + k_t^{-1} fe_{t|t-1} \tag{2}$$

The issue seems to be: if $fe_{t|t-1} \approx 0$, then the gain is irrelevant for learning because $fe_{t|t-1}$ figures into both equations. So the idea is to decouple the two equations by changing the timing of one of the forecast errors. Note:

$$\begin{aligned} fe_{t|t-1} &= \pi_t - (\bar{\pi}_{t-1} + b s_{t-1}) && (3) \\ &= \pi_t - \bar{\pi}_{t-1} \qquad \text{since shocks iid and } b \text{ is the RE transition matrix} && (4) \end{aligned}$$

So what I can try is to use an older forecast error in equation (2). Try $fe_{t|t-1} \equiv \pi_t - \bar{\pi}_{t-2}$.

**Figure 4:** Estimates for $N = 100$, changing the forecast error timing in the updating equation, imposing convexity with weight 100K, truth with $nfe = 5, fe \in (-2, 2)$



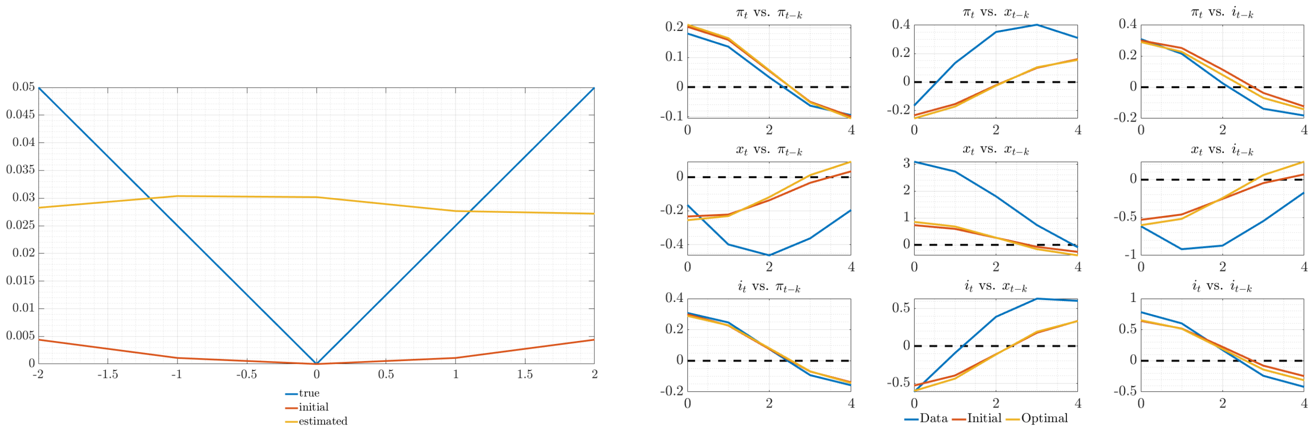

(a) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$

(b) Autocovariogram

A little more symmetric, but no dramatic improvement.

## Point #+1: do $N$ simulations instead of $N$ estimations

**Figure 5:** Estimates for $N = 100$, targeting mean moments in a single estimation instead of $N$ estimations of individual moments, imposing convexity with weight 100K, truth with $nfe = 5, fe \in (-2, 2)$




(a) $\alpha^{true}, \alpha_0, mean(\hat{\alpha})$

(b) Autocovariogram

The difference is striking!

## Point #+2: introduce expectations series

# A Model summary

$$x_t = -\sigma i_t + \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} \beta^{T-t}\big((1-\beta)x_{T+1} - \sigma(\beta i_{T+1} - \pi_{T+1}) + \sigma r_T^n\big) \tag{A.1}$$

$$\pi_t = \kappa x_t + \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t}\big(\kappa\alpha\beta x_{T+1} + (1-\alpha)\beta\pi_{T+1} + u_T\big) \tag{A.2}$$

$$i_t = \psi_\pi \pi_t + \psi_x x_t + \bar{i}_t \qquad \text{(if imposed)} \tag{A.3}$$

$$\text{PLM:} \qquad \hat{\mathbb{E}}_t z_{t+h} = a_{t-1} + b h_x^{h-1} s_t \quad \forall h \geq 1 \qquad b = g_x\, h_x \tag{A.4}$$

$$\text{Updating:} \qquad a_t = a_{t-1} + k_t^{-1}\big(z_t - (a_{t-1} + b s_{t-1})\big) \tag{A.5}$$

$$\text{Anchoring function:} \qquad k_t^{-1} = \rho_k k_{t-1}^{-1} + \gamma_k f e_{t-1}^2 \tag{A.6}$$

$$\text{Forecast error:} \qquad fe_{t-1} = z_t - (a_{t-1} + b s_{t-1}) \tag{A.7}$$

$$\text{LH expectations:} \qquad f_a(t) = \frac{1}{1-\alpha\beta} a_{t-1} + b(\mathbb{I}_{nx} - \alpha\beta h)^{-1} s_t \qquad f_b(t) = \frac{1}{1-\beta} a_{t-1} + b(\mathbb{I}_{nx} - \beta h)^{-1} s_t \tag{A.8}$$

This notation captures vector learning ($z$ learned) for intercept only. For scalar learning, $a_t = \begin{pmatrix}\bar{\pi}_t & 0 & 0\end{pmatrix}'$ and $b_1$ designates the first row of $b$. The observables $(\pi, x)$ are determined as:

$$x_t = -\sigma i_t + \begin{bmatrix}\sigma & 1-\beta & -\sigma\beta\end{bmatrix} f_b + \sigma \begin{bmatrix}1 & 0 & 0\end{bmatrix} (\mathbb{I}_{nx} - \beta h_x)^{-1} s_t \tag{A.9}$$

$$\pi_t = \kappa x_t + \begin{bmatrix}(1-\alpha)\beta & \kappa\alpha\beta & 0\end{bmatrix} f_a + \begin{bmatrix}0 & 0 & 1\end{bmatrix} (\mathbb{I}_{nx} - \alpha\beta h_x)^{-1} s_t \tag{A.10}$$

# B Target criterion

The target criterion in the simplified model (scalar learning of inflation intercept only, $k_t^{-1} = \mathbf{g}(fe_{t-1})$):

$$\pi_t = -\frac{\lambda_x}{\kappa}\Big\{ x_t - \frac{(1-\alpha)\beta}{1-\alpha\beta}\Big(k_t^{-1} + ((\pi_t - \bar{\pi}_{t-1} - b_1 s_{t-1}))\mathbf{g}_\pi(t)\Big)$$
$$\Big(\mathbb{E}_t \sum_{i=1}^{\infty} x_{t+i} \prod_{j=0}^{i-1}(1 - k_{t+1+j}^{-1} - (\pi_{t+1+j} - \bar{\pi}_{t+j} - b_1 s_{t+j})\mathbf{g}_{\bar{\pi}}(t+j))\Big)\Big\} \tag{B.1}$$

where I'm using the notation that $\prod_{j=0}^{0} \equiv 1$. For interpretation purposes, let me rewrite this as follows:

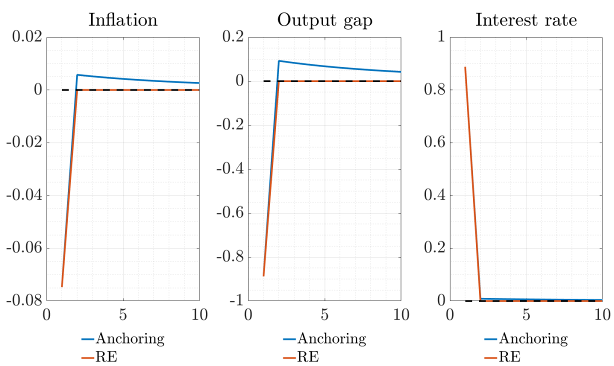
$$\pi_t = -\frac{\lambda_x}{\kappa} x_t + \frac{\lambda_x}{\kappa}\frac{(1-\alpha)\beta}{1-\alpha\beta}\Big(k_t^{-1} + fe_{t|t-1}^{eve}\mathbf{g}_\pi(t)\Big)\mathbb{E}_t \sum_{i=1}^{\infty} x_{t+i}$$
$$-\frac{\lambda_x}{\kappa}\frac{(1-\alpha)\beta}{1-\alpha\beta}\Big(k_t^{-1} + fe_{t|t-1}^{eve}\mathbf{g}_\pi(t)\Big)\Big(\mathbb{E}_t \sum_{i=1}^{\infty} x_{t+i} \prod_{j=0}^{i-1}(k_{t+1+j}^{-1} + fe_{t+1+j|t+j}^{eve})\mathbf{g}_{\bar{\pi}}(t+j)\Big) \tag{B.2}$$

Interpretation: tradeoffs from discretion in RE + effect of current level and change of the gain on future tradeoffs + effect of future expected levels and changes of the gain on future tradeoffs

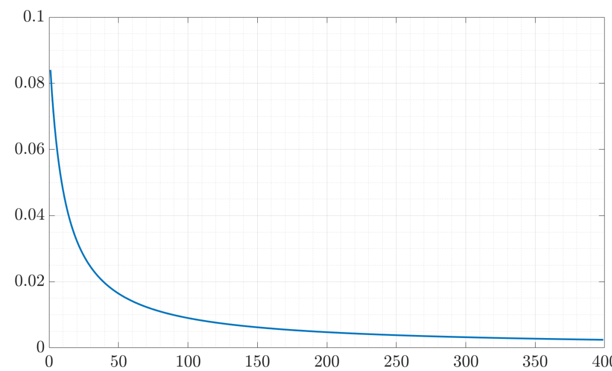
# C Impulse responses to iid monpol shocks across a wide range of learning models

$T = 400, N = 100, n_{drop} = 5$, shock imposed at $t = 25$, calibration as above, Taylor rule assumed to be known, PLM = learn constant only, of inflation only.

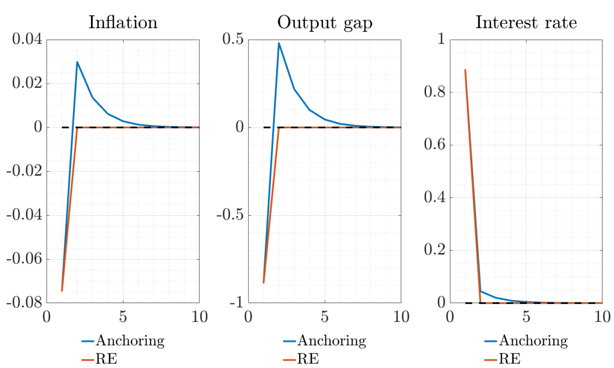
**Figure 6:** IRFs and gain history (sample means)



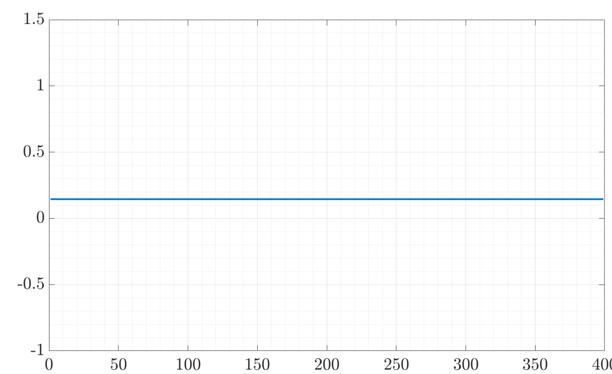

(a) Decreasing gain learning



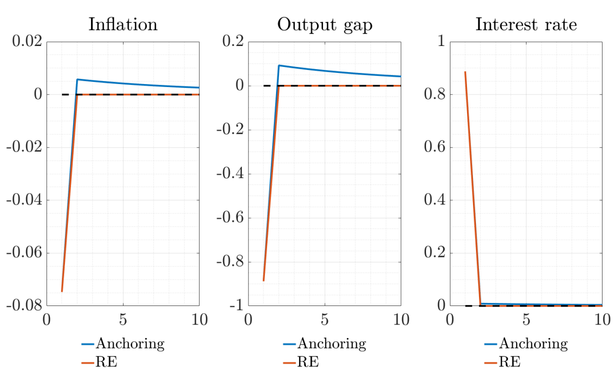
(b) Mean gain



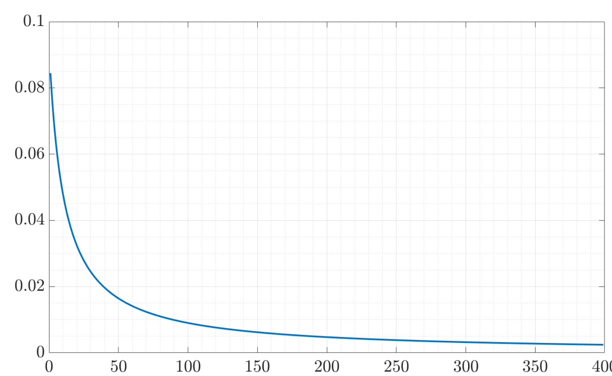

(c) Constant gain learning



(d) Mean gain

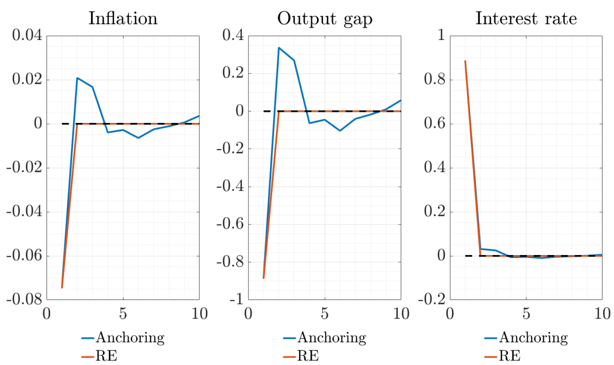



(e) CEMP criterion (vector)

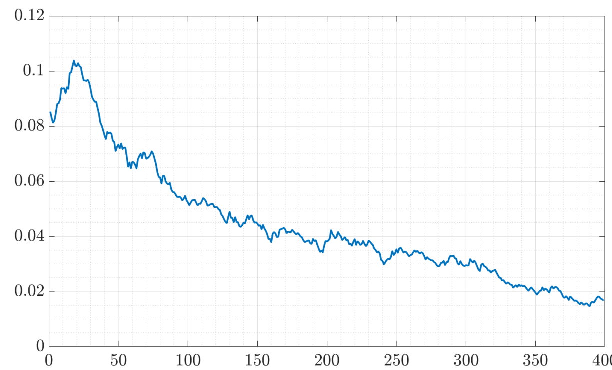


(f) Mean gain

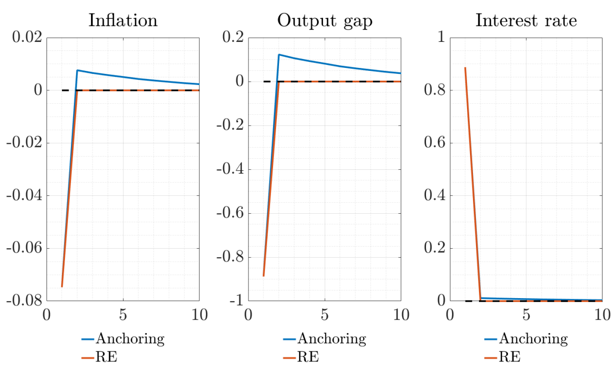
**Figure 7:** IRFs and gain history (sample means), continued



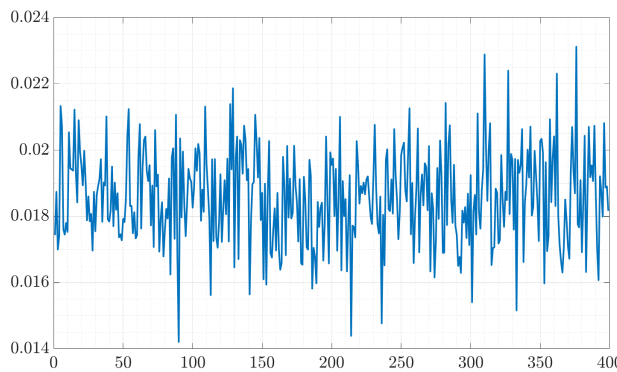

(a) CUSUM criterion (vector)



(b) Mean gain




(c) Smooth criterion, approximated, using $\alpha^{true} = (0.05; 0.025; 0; 0.025; 0.05)$, on $fe \in (-2, 2)$.



(d) Mean gain