Online Heavy-tailed Change-point detection (Supplementary Materials)

Abishek Sankararaman[*1] **Balakrishnan (Murali) Narayanaswamy**[1]

[1]AWS AI Labs

A CHANGE-POINT LOCALIZATION

Algorithm 1 Online `Clipped-SGD` Change Point Detection and Localization

Input: $(\eta_t)_{t \geq 1}$, $\lambda > 0$, $\theta_0 \in \Theta$, $\delta \in (0,1)$ FPR guarantee
$r \leftarrow 1$
$\widehat{\theta}_{t,t-1} \leftarrow \theta_0$, for all $t \geq 1$.
Set $\tau_c^{(0)} \leftarrow 0$
Set `Num-change-points` $\leftarrow 0$
for each time $t = 1, 2, \cdots,$ **do**
  Receive sample X_t
  $\widehat{\theta}_{s,t} \leftarrow \prod_{\theta}(\widehat{\theta}_{s,t-1} - \eta_{t-s}\text{clip}(X_t - \widehat{\theta}_{s,t-1}, \lambda))$, for every $r \leq s \leq t$.
  if $\exists s \in (r,t)$ such that $\|\widehat{\theta}_{r:s} - \widehat{\theta}_{s+1:t}\|_2^2 > \mathcal{B}\left(s-r, \frac{\delta}{2(t-r)(t-r+1)}\right) + \mathcal{B}\left(t-s-1, \frac{\delta}{2(t-r)(t-r+1)}\right)$ $\{B(\cdot,\cdot)$ is defined in Equation (5)$\}$ **then**
    Set **Restart**$_t \leftarrow 1$ {Change point detected}
    Set `Num-change-points` $\leftarrow$`Num-change-points` $+1$ {Increment number of change-points detected}
    Output time interval $[\inf\{s \in (r,t) \text{ s.t. } \mathfrak{B}(r,s,t,\delta) = 1\}, \sup\{s \in (r,t) \text{ s.t. } \mathfrak{B}(r,s,t,\delta) = 1\}]$ as the location of the change-point $\{\mathfrak{B}()$ defined in Equation (8)$\}$
    $r \leftarrow t + 1$
  else
    Set **Restart**$_t \leftarrow 0$
  end if
end for

B PROOF FOR ROBUST ESTIMATION IN THEOREM 3.1

We follow the same proof architecture as that of Proof of [Tsai et al., 2022].

Fix a time $t \in \mathbb{N}$. We define a sequence of random variable $(\psi_t)_{t \geq 1}$ as follows.

$$\psi_t := \text{clip}((X_t - \widehat{\theta}_{t-1}), \lambda) - (\theta^* - \widehat{\theta}_{t-1}),$$

Consider any time t. We have

$$\|\theta_t - \theta^*\|_2^2 = \|\prod_{\Theta}(\widehat{\theta}_{t-1} - \eta_t \text{clip}(X_t - \widehat{\theta}_{t-1}, \lambda)) - \theta^*\|_2^2, \tag{1}$$

$$\overset{(a)}{\leq} \|\widehat{\theta}_{t-1} - \eta_t \text{clip}(X_t - \widehat{\theta}_{t-1}, \lambda) - \theta^*\|_2^2, \tag{2}$$

Accepted for the 39th Conference on Uncertainty in Artificial Intelligence (UAI 2023).

$$
\begin{aligned}
&= \|\widehat{\theta}_{t-1} - \eta_t(\psi_t + (\theta^* - \widehat{\theta}_{t-1})) - \theta^*\|_2^2, \\
&= \|\widehat{\theta}_{t-1} - \theta^*\|_2^2 + \eta_t^2\|\psi_t + (\theta^* - \widehat{\theta}_{t-1})\|_2^2 - 2\eta_t\langle\widehat{\theta}_{t-1} - \theta^*, \psi_t + (\theta^* - \widehat{\theta}_{t-1})\rangle, \\
&\overset{(b)}{\leq} \|\widehat{\theta}_{t-1} - \theta^*\|_2^2 + 2\eta_t^2\|\psi_t\|_2^2 + 2\eta_t^2\|(\theta^* - \widehat{\theta}_{t-1})\|_2^2 - 2\eta_t\langle\widehat{\theta}_{t-1} - \theta^*, \psi_t + (\theta^* - \widehat{\theta}_{t-1})\rangle, \qquad (3)
\end{aligned}
$$

Step (a) follows since Θ is a convex set, $\|\mathcal{P}_\Theta(\widehat{\theta}_t) - \theta^*\| \leq \|\widehat{\theta}_t - \theta^*\|$, since $\theta^* \in \Theta$. In step (b), we use the fact that $\|a+b\|_2^2 \leq 2\|a\|_2^2 + 2\|b\|_2^2$, for all $a, b \in \mathbb{R}^d$. Substituting Equation (35) into (3), we get that

$$
\begin{aligned}
\|\theta^* - \theta_t\|_2^2 \leq \|\widehat{\theta}_{t-1} - \theta^*\|_2^2 + 2\eta_t^2\|\psi_t\|_2^2 - 2\eta_t\langle\widehat{\theta}_{t-1} - \theta_t^*, \psi_t\rangle \\
+ 2\eta_t^2\Big((M+m)\langle(\theta^* - \widehat{\theta}_{t-1}), \widehat{\theta}_{t-1} - \theta_t^*\rangle - mM\|\widehat{\theta}_{t-1} - \theta^*\|_2^2\Big) - 2\eta_t\langle(\theta^* - \widehat{\theta}_{t-1}), \widehat{\theta}_{t-1} - \theta_t^*\rangle.
\end{aligned}
$$

Re-arranging the equation above yields

$$
\begin{aligned}
\|\theta^* - \theta_t\|_2^2 \leq (1 - 2\eta_t^2 mM)\|\widehat{\theta}_{t-1} - \theta^*\|_2^2 + 2\eta_t^2\|\psi_t\|_2^2 - 2\eta_t\langle\widehat{\theta}_{t-1} - \theta^*, \psi_t\rangle \\
- 2\eta_t(1 - \eta_t\left((M+m)\right)\langle(\theta^* - \widehat{\theta}_{t-1}), \widehat{\theta}_{t-1} - \theta^*\rangle.
\end{aligned}
$$

Further substituting Equation (34) into the display above yields that

$$
\begin{aligned}
\|\theta^* - \widehat{\theta}_t\|_2^2 &\leq (1 - 2\eta_t m + 2\eta_t^2 m^2)\|\widehat{\theta}_{t-1} - \theta^*\|_2^2 + 2\eta_t^2\|\psi_t\|_2^2 - 2\eta_t\langle\widehat{\theta}_{t-1} - \theta^*, \psi_t\rangle, \\
&\leq (1 - \eta_t m)\|\widehat{\theta}_{t-1} - \theta^*\|_2^2 + 2\eta_t^2\|\psi_t\|_2^2 - 2\eta_t\langle\widehat{\theta}_{t-1} - \theta^*, \psi_t\rangle,
\end{aligned}
$$

where the inequality comes from the fact that if $\eta_t m < 1 \implies 2\eta_t m - 2\eta_t^2 m^2 > \eta m$.

$$
\|\theta^* - \widehat{\theta}_t\|_2^2 \leq (1 - \eta_t m)\|\widehat{\theta}_{t-1} - \theta^*\|_2^2 + 2\eta_t^2\|\psi_t\|_2^2 - 2\eta_t\langle\widehat{\theta}_{t-1} - \theta^*, \psi_t\rangle. \qquad (4)
$$

Unrolling the recursion yields,

$$
\begin{aligned}
\|\theta^* - \widehat{\theta}_t\|_2^2 \leq \prod_{u=1}^{t}(1 - \eta_u m)\|\theta_1 - \theta^*\|_2^2 + 2\eta_t^2\sum_{s=1}^{t-1}\prod_{u=1}^{s}(1 - \eta_{t-u+1}m)\|\psi_{t-s+1}\|_2^2 \\
-2\eta_t\sum_{s=1}^{t-1}\prod_{u=1}^{s}(1 - \eta_{t-u+1}m)\langle\theta_{t-s} - \theta^*, \psi_{t-s+1}\rangle.
\end{aligned}
$$

Using the fact that $\prod_{u=1}^{s}(1 - \eta_{t-u+1}m) = \frac{(t-s+\gamma-3)(t-s+\gamma-2)}{(t+\gamma)(t+\gamma-1)}$, we get that

$$
\|\theta^* - \widehat{\theta}_t\|_2^2 \leq \frac{(\gamma-2)(\gamma-1)\|\theta_1 - \theta^*\|_2^2}{(t+\gamma)(t+\gamma-1)} \qquad (5)
$$

$$
-2\eta_t\sum_{s=1}^{t-1}\frac{(t-s+\gamma-3)(t-s+\gamma-2)\langle\theta_{t-s} - \theta^*, \psi_{t-s+1}\rangle}{(t+\gamma)(t+\gamma-1)}. \qquad (6)
$$

Denote by $\psi_t := \psi_t^{(b)} + \psi_t^{(v)}$, where $\psi_t^{(b)} := \mathbb{E}_{Z_t}[\psi_t|\mathcal{F}_{t-1}]$ and $\psi_t^{(v)} := \psi_t - \psi_t^{(b)}$. Using this in the display above and using that fact that $\|a+b\|_2^2 \leq 2\|a\|_2^2 + 2\|b\|_2^2$, we get

$$
\begin{aligned}
\|\theta_t^* - \theta\|_2^2 \leq \frac{(\gamma-1)(\gamma-2)\|\theta_1 - \theta^*\|_2^2}{(t+\gamma)(t+\gamma-1)} + 4\eta_t^2\sum_{s=1}^{t-1}\frac{(t-s+\gamma-3)(t-s+\gamma-2)\|\psi_{t-s+1}\|_2^2}{(t+\gamma)(t+\gamma-1)} + \\
- 2\eta_t\sum_{s=1}^{t-1}\frac{(t-s+\gamma-3)(t-s+\gamma-2)\langle\theta_{t-s} - \theta^*, \psi_{t-s+1}^{(b)}\rangle}{(t+\gamma)(t+\gamma-1)} \\
- 2\eta_t\sum_{s=1}^{t-1}\frac{(t-s+\gamma-3)(t-s+\gamma-2)\langle\theta_{t-s} - \theta^*, \psi_{t-s+1}^{(v)}\rangle}{(t+\gamma)(t+\gamma-1)}.
\end{aligned}
$$

Further simplifying by adding and subtracting $\mathbb{E}_{Z_t}[\|\psi_t^{(v)}\|_2^2|\mathcal{F}_{t-1}]$ to be above display, we get

$$\|\theta^* - \widehat{\theta}_t\|_2^2 \leq \frac{(\gamma-1)(\gamma-2)\|\theta_1 - \theta^*\|_2^2}{(t+\gamma)(t+\gamma-1)} + 4\eta_t^2 \sum_{s=1}^{t-1} \frac{(t-s+\gamma-3)(t-s+\gamma-2)\|\psi_{t-s+1}^{(b)}\|_2^2}{(t+\gamma)(t+\gamma-1)} \tag{7}$$

$$+ 4\eta_t^2 \sum_{s=1}^{t-1} \frac{(t-s+\gamma-3)(t-s+\gamma-2)\mathbb{E}_{Z_{t-s+1}}[\|\psi_{t-s+1}^{(v)}\|_2^2|\mathcal{F}_{t-s}]}{(t+\gamma)(t+\gamma-1)}$$

$$+ 4\eta_t^2 \sum_{s=1}^{t-1} \frac{(t-s+\gamma-3)(t-s+\gamma-2)(\|\psi_{t-s+1}^{(v)}\|_2^2 - \mathbb{E}_{Z_{t-s+1}}[\|\psi_{t-s+1}^{(v)}\|_2^2|\mathcal{F}_{t-s}])}{(t+\gamma)(t+\gamma-1)} \tag{8}$$

$$- 2\eta_t \sum_{s=1}^{t-1} \frac{(t-s+\gamma-3)(t-s+\gamma-2)\langle \theta_{t-s} - \theta^*, \psi_{t-s+1}^{(b)}\rangle}{(t+\gamma)(t+\gamma-1)}$$

$$- 2\eta \sum_{s=1}^{t-1} \frac{(t-s+\gamma-3)(t-s+\gamma-2)\langle \theta_{t-s} - \theta^*, \psi_{t-s+1}^{(v)}\rangle}{(t+\gamma)(t+\gamma-1)}. \tag{9}$$

Lemma B.1 (Lemma F.5 [Gorbunov et al., 2020]). *If $\lambda \geq 2G$, the following inequalities hold almost-surely for all times t.*

$$\|\psi_t^{(v)}\| \leq 2\lambda \mathbf{1}_{\sigma>0} \tag{10}$$

$$\|\psi_t^{(b)}\|_2 \leq \frac{4\sigma^2}{\lambda} \tag{11}$$

$$\mathbb{E}_{Z_t}[\|\psi_t^{(v)}\|_2^2|\mathcal{F}_{t-1}] \leq 10\sigma^2 \tag{12}$$

Simplifying Equation (9) using bounds in Lemma B.1, along with the fact that for all $1 \leq s \leq t$ and $\gamma \geq 1$, $\frac{(t-s+\gamma-3)(t-s+\gamma-2)}{(t+\gamma)(t+\gamma-1)} \leq \frac{t-s+\gamma}{t+\gamma}$ we get

$$\|\theta^* - \widehat{\theta}_t\|_2^2 \leq \frac{(\gamma-1)(\gamma-2)\|\theta_1 - \theta^*\|_2^2}{(t+\gamma)(t+\gamma-1)} + \frac{16\eta_t^2\sigma^2}{\lambda} \sum_{s=1}^{t-1} \frac{t-s+\gamma}{t+\gamma} + 4\eta_t^2\sigma^2 \sum_{s=1}^{t-1} \frac{t-s+\gamma}{t+\gamma}$$

$$+ 4\eta_t^2 \sum_{s=1}^{t-1} \frac{(t-s+\gamma)(\|\psi_{t-s+1}^{(v)}\|_2^2 - \mathbb{E}_{Z_{t-s+1}}[\|\psi_{t-s+1}^{(v)}\|_2^2|\mathcal{F}_{t-s+1}])}{t+\gamma}$$

$$+ 2\eta_t \sum_{s=1}^{t-1} \frac{(t-s+\gamma)\|\theta_{t-s} - \theta^*\|\|\psi_{t-s+1}^{(b)}\|}{t+\gamma} + -2\eta_t \sum_{s=1}^{t-1} \frac{(t-s+\gamma)\langle \theta_{t-s} - \theta^*, \psi_{t-s+1}^{(v)}\rangle}{t+\gamma}. \tag{13}$$

Further applying the bound that $\|\psi_t^{(b)}\| \leq \frac{4\sigma^2}{\lambda}$

$$\|\theta^* - \widehat{\theta}_t\|_2^2 \leq \frac{(\gamma-1)(\gamma-2)\|\theta_1 - \theta^*\|_2^2}{(t+\gamma)(t+\gamma-1)} + \underbrace{\left(\frac{16\eta_t^2\sigma^2}{\lambda} + 4\eta_t^2\sigma^2\right) \sum_{s=1}^{t-1} \frac{t-s+1}{t+\gamma}}_{\text{Term 1}}$$

$$+ \underbrace{4\eta_t^2 \sum_{s=1}^{t-1} \frac{(t-s+\gamma)(\|\psi_{t-s+1}^{(v)}\|_2^2 - \mathbb{E}_{Z_{t-s+1}}[\|\psi_{t-s+1}^{(v)}\|_2^2|\mathcal{F}_{t-s+1}])}{t+\gamma}}_{\text{Term 2}}$$

$$+ \underbrace{\frac{8\sigma^2\eta_t}{\lambda} \sum_{s=1}^{t-1} \frac{(t-s+\gamma)\|\theta_{t-s} - \theta^*\|}{t+\gamma}}_{\text{Term 3}} \underbrace{-2\eta_t \sum_{s=1}^{t-1} \frac{(t-s+\gamma)\langle \theta_{t-s} - \theta^*, \psi_{t-s+1}^{(v)}\rangle}{t+\gamma}}_{\text{Term 4}}. \tag{14}$$

B.1 PROBABILISTIC ANALYSIS

Definitions

For every $t \geq 1$, denote by the constant

$$C_t = \max\left(\frac{1024\sigma^4}{G^2 m^2 \lambda^2}, \frac{8\lambda\sqrt{\ln\left(\frac{2t^3}{\delta}\right)}}{\gamma^2 G}\right). \tag{15}$$

Denote by the deterministic constant $\xi_u^{(t)}$ for $u = 1, \cdots, t$ as

$$\left(\xi_u^{(t)}\right)^2 := C_t\Bigg[\left(\frac{16\sigma^2}{\lambda} + 4\sigma^2\right)\frac{1}{2m^2(u+1)} + \frac{96\lambda^2 \ln\left(\frac{2t^3}{\delta}\right)\sigma(\sigma+1)}{m(u+\gamma)\sqrt{u+1}}\Bigg]. \tag{16}$$

From the definition, the following in-equalities hold.

Proposition B.2. *For all times $u \in \{1, \cdots, t\}$,*

$$\sum_{s=1}^{u-1}(u-s+\gamma)\xi_s^{(t)} \leq 2(u+\gamma)\sqrt{u+1}\xi_u^{(t)}, \tag{17}$$

$$\sum_{s=1}^{u-1}(\xi_s^{(t)})^2 \leq 2(u+1)\ln(u+1)(\xi_u^{(t)})^2 \tag{18}$$

Proof. This follows from the following fact.

Proposition B.3. *For all $u \in \mathbb{N}$ and $\gamma \geq 0$, we have*

$$\sum_{s=1}^{u-1}\frac{u-s+\gamma}{\sqrt{u+1}} \leq 2(u+\gamma)\sqrt{u+1}.$$

□

For each time $u \in \{1, \cdots, t\}$, denote by the random variable $\nu_u^{(t)}$ by

$$\nu_u^{(t)} := \begin{cases} \theta_u - \theta^* & \text{if } \|\theta_u - \theta^*\|^2 \leq (\xi_u^{(t)})^2 + \frac{C_t\gamma^2 G^2}{(u+1)} \\ 0 & \text{if otherwise} \end{cases}$$

For every $u \in \{1, \cdots, t\}$, denote by the event $\mathcal{E}_{u;1}^{(t)}$ to be the one in which the following inequality holds for all $u \in \{1, \cdots, t\}$.

$$\mathcal{E}_{u;1}^{(t)} := \Bigg\{4\eta_t^2 \sum_{s=1}^{u-1}\frac{(u-s+\gamma)(\|\psi_{u-s+1}^{(v)}\|_2^2 - \mathbb{E}_{Z_{u\ \ s+1}}[\|\psi_{u-s+1}^{(v)}\|_2^2|\mathcal{F}_{u-s+1}])}{t+\gamma} \leq \frac{96\lambda^2\ln\left(\frac{2t^2(t+1)}{\delta}\right)\sigma(\sigma+1)}{m(u+\gamma)\sqrt{u+1}}\Bigg\}. \tag{19}$$

and $\mathcal{E}_{u;2}^{(t)}$ as

$$\mathcal{E}_{u;2}^{(t)} := \Bigg\{-2\eta_u\sum_{s=1}^{u-1}\frac{(u-s+\gamma)\langle \upsilon_{u-s}, \psi_{u-s+1}^{(v)}\rangle}{t+\gamma} \leq \frac{\xi_u^{(t)}\ln\left(\frac{2t^2(t+1)}{\delta}\right)}{10\sqrt{u+1}} + \frac{C_u\gamma^2G^2}{4(u+1)}.\Bigg\} \tag{20}$$

Denote by the event $\mathcal{E}^{(t)}$ as

$$\mathcal{E}^{(t)} := \bigcap_{u=1}^{t}\left(\mathcal{E}_{u;1}^{(t)} \cap \mathcal{E}_{u;2}^{(t)}\right). \tag{21}$$

Lemma B.4. For all $t \quad 1$,

$$P[E^{(t)}] \quad 1 \quad \frac{}{t(t+1)}:$$

We now prove by induction hypothesis that

Lemma B.5. For every t, under the event $E^{(t)}$, the following holds.

$$k b_u \qquad k_2^2 \quad \frac{C_t \;\; {}^2 G^2}{(u+1)^{\;2}} + (\;\; {}_u^{(t)})^2; \tag{22}$$

for all $u \; 2 \; f \; 1; \quad ; t g$.

Proof. Proof of Lemma B.1 We will prove this lemma by induction on u by analyzing Equation (14). The base-case of $u = 1$ holds trivially with probability 1 since $C_t > 1, 8t \quad 1$ and $\; > \; 2$.

Now, assume that on the event $E^{(t)}$, the induction hypothesis in Equation (22) holds for all times $1; \quad ; u \quad 1$. We prove this by expanding Equation (14) and bounding each of the terms.

Term 1

It is easy to verify that

$$\frac{16 \;\; {}_u^2 \;\; {}^2}{} + 4 \;\; {}_u^2 \;\; {}^2 \sum_{s=1}^{u \; 1} \frac{u \quad s+}{u+} \qquad \frac{16 \;\; {}^2}{} + 4 \;\; {}^2 \quad \frac{u}{2m^2(u+ \quad)^2};$$

$$\frac{\frac{16 \;\; {}^2}{} + 4 \;\; {}^2}{2m^2(u+1)}:$$

The last inequality follows since $^2 > 1$.

Term 2

First notice that

$$4 \;\; {}_u^2 \sum_{s=1}^{u \; 1} \frac{(u \quad s+ \quad)(k \;\; {}_{u \;\; s+1}^{(v)} k_2^2 \quad E_{Z_{u \;\; s+1}} [k \;\; {}_{u \;\; s+1}^{(v)} k_2^2 j F_{u \;\; s+1}])}{t+}$$

$$\frac{4 \;\; {}_u}{u+} \sum_{s=1}^{u \; 1} (k \;\; {}_{u \;\; s+1}^{(v)} k_2^2 \quad E_{Z_{u \;\; s+1}} [k \;\; {}_{u \;\; s+1}^{(v)} k_2^2 j F_{u \;\; s+1}])$$

From the de nition of event $E^{(t)}$ in Equation (21), we get that

$$\text{Term}\,2 \quad \frac{96 \;\; {}^2 \ln \frac{2t^2(t+1)}{} \qquad (\quad +1)}{m(u+ \quad) \sqrt{u+1}}:$$

Term 3

$$\frac{8 \;\; {}^2 \;\; {}_u}{} \sum_{s=1}^{u \; 1} \frac{(u \quad s+ \quad) k \;\; {}_{u \;\; s} \qquad k}{u+} \qquad \frac{8 \;\; {}^2}{m \;\; (u+ \quad)^2} \sum_{s=1}^{u \; 1} \;\; (u \quad s+ \quad) \;\; {}_{u \;\; s}^{(t)} + \sqrt{\overline{C_t}} \; G \; \frac{(u \quad s+ \quad)}{(u \quad s+1)} \quad ;$$

$$\overset{(18)}{} \frac{16\ {}^2 \mathrm{p}\ \overline{(u+1)}\ {}^{(t)}_{u}}{m(u+\ \ )} + \frac{8 \mathrm{p}\ \overline{C_t}\ \ {}^2\ {}^2 G u}{m\ \ (u+\ \ )^2};$$

$$\frac{16\ {}^2 \mathrm{p}\ \overline{(u+1)}\ {}^{(t)}_{u}}{m(u+\ \ )} + \frac{8 \mathrm{p}\ \overline{C_t}\ \ {}^2\ {}^2 G}{m\ \ (u+\ \ )};$$

$$\overset{(a)}{} \frac{{}^{(t)}_{u}}{10 \mathrm{p}\ \overline{u+1}} + \frac{C_t\ \ {}^2 G^2}{4(u+1)}:$$

The last inequality follows since $\frac{320\ \ ^2}{m} + 1 \quad =) \quad \frac{8\ \ ^2(u+1)^{1=2} \log(u+1)}{m(u+\ \ )} \quad \frac{1}{10 \mathrm{p}\ \overline{u+1}}$, for all $u \quad t$ and the fact that $C_t \quad \frac{1024\ \ ^4}{G^2 m^2\ \ ^2}$.

Term 4

The de nition of event $E^{(t)}$ in Equation (21) gives that Term 4 $\quad \frac{{}^{(t)}_{u} \ln \frac{2t^2(t+1)}{}}{10 \mathrm{p}\ \overline{u+1}} + \frac{C_t\ \ ^2 G^2}{4(u+1)}$

Now, adding in the bounds together into Equation (14),

$$k b_u \quad k_2^2 \quad \frac{^2 G^2}{u+1} + \frac{\frac{16\ \ ^2}{} + 4\ \ ^2}{2m^2(u+1)} + \frac{{}^{(t)}_{u}}{10 \mathrm{p}\ \overline{u+1}} + \frac{1600\ ^2 \ln \frac{2t^2(t+1)}{} \quad (\ \ +1)}{m(u+\ \ ) \mathrm{p}\ \overline{u+1}}$$

$$+ \frac{{}^{(t)}_{u} \ln \frac{2t^2(t+1)}{}}{10 \mathrm{p}\ \overline{u+1}} + \frac{C_t\ \ ^2 G^2}{2(u+1)}:$$

Now using the fact that $\frac{{}^{(t)}_{u} \ln \frac{2t^3}{}}{\mathrm{p}\ \overline{u+1}} \quad ({}^{(t)}_{u})^2$, we get that

$$k b_u \quad k_2^2 \quad 1 + \frac{C_t}{2} \quad \frac{^2 G^2}{u+1} + \frac{\frac{16\ \ ^2}{} + 4\ \ ^2}{2m^2(u+1)} + \frac{({}^{(t)}_{u})^2}{5} + \frac{96\ ^2 \ln \frac{2t^2(t+1)}{} \quad (\ \ +1)}{m(u+\ \ ) \mathrm{p}\ \overline{u+1}}:$$

Substituting the de nition of ${}^{(t)}_{u}$ from Equation (16), we get that

$$k b_u \quad k_2^2 \quad 1 + \frac{C_t}{2} \quad \left[\frac{^2 G^2}{u+1} + \frac{\frac{16\ \ ^2}{} + 4\ \ ^2}{2m^2(u+1)} + \frac{96\ ^2 \ln \frac{2t^2(t+1)}{} \quad (\ \ +1)}{m(u+\ \ ) \mathrm{p}\ \overline{u+1}} \right];$$

$$({}^{(t)}_{u})^2 + \frac{C_t\ \ ^2 G^2}{u+1}:$$

The last inequality follows since $C_t = \max \left(\frac{1024\ \ ^4}{G^2 m^2\ \ ^2}; \frac{8 \sqrt{\ln \frac{2t^3}{}}}{^2 G} \right) \quad =) \quad C_t \quad 2.$

□

□

B.2 PROOF OF LEMMA B.4

We rst reproduce an useful result.

Lemma B.6 (Freedman's inequality [Victor, 1999]) Suppose $Y_1; \quad ; Y_T$ is a bounded martingale with respect to a ltration $(F_t)_{t=0}^T$ with $E[Y_t j F_{t\ \ 1}] = 0$ and $P[jY_t j \quad B] = 1$ for all $t\ 2\ f\ 1; \quad ; T g$. Denote by $V_s := \mathrm{P}_{n=1}^{s} \mathrm{Var}(Y_n j F_{n\ \ 1})$ be the sum of conditional variances. Then, for every $a; v > 0$,

$$P \quad 9n\ 2\ [1; T] \text{ such that } \sum_{t=1}^{n} Y_t \quad a \text{ and } V_n \quad v \quad ! \quad \exp \quad \frac{a^2}{2(v + Ba)} \quad : \qquad (23)$$

Re-arranging the above inequality, we see that if

$$a \quad B\ln\frac{2T}{} + \sqrt{\left(B\ln\frac{2T}{}\right)^2 + 2v\ln\frac{2T}{}}; \qquad (24)$$

then the RHS of Equation (23) is bounded above by $\frac{}{2}$

Proof of Lemma B.4. Proof of Equation (19)

Fix a $u \in \{1; \ldots; t\}$. For $s \in \{1; \ldots; u-1\}$, denote by the random variable $Y_s^{(u)} := \frac{(u-s+\;)}{u+\;}(\|{}_{u-s+1}^{(v)}\|_2^2 - E_{Z_{u-s+1}}[\|{}_{u-s+1}^{(v)}\|_2^2 | F_{u-s}])$. Thus,

$$4\,{}_u^2 \sum_{s=1}^{-1} \frac{(u-s+\;)(\|{}_{u-s+1}^{(v)}\|_2^2 - E_{Z_{u-s+1}}[\|{}_{u-s+1}^{(v)}\|_2^2 | F_{u-s+1}])}{u+\;} \quad 4\,{}_u^2 \sum_{s=1}^{-1} Y_s^{(u)}:$$

Observe that the sequence $(Y_s^{(u)})_{s=1}^{u-1}$ is a martingale difference sequence with respect to the filtration $(G_s)_{s=1}^{t-1}$, where $G_s := F_{u-s}$. Furthermore, observe that with probability 1, $|Y_s^{(u)}| \quad 4\;{}^2 1_{>0} + 4\;{}^2 1_{>0} \quad 8\;{}^2 1_{>0}$. We can bound the conditional variance as

$$\begin{aligned}
\sum_{s=1}^{-1} \mathrm{Var}(Y_s^{(u)}|G_s) &\quad \sum_{s=1}^{-1}\left(\frac{(u-s+\;)}{u+\;}\right)^2 E_{Z_{u-s}}[(\|{}_{u-s+1}^{(v)}\|_2^2 - E_{Z_{u-s+1}}[\|{}_{u-s+1}^{(v)}\|_2^2|F_{u-s}])^2|F_{u-s}];\\
&\overset{10}{\quad} 8\;{}^2 \sum_{s=1}^{-1} E_{Z_{u-s}}[|\|{}_{u-s+1}^{(v)}\|_2^2 - E_{Z_{u-s+1}}[\|{}_{u-s+1}^{(v)}\|_2^2|F_{u-s}]||F_{u-s}];\\
&\quad 8\;{}^2 \sum_{s=1}^{-1} 2E_{Z_{u-s}}[\|{}_{u-s+1}^{(v)}\|_2^2|F_{u-s}];\\
&\overset{12}{\quad} 160\;{}^2\;{}^2(u-1):
\end{aligned}$$

Now, putting $B := 8\;{}^2$ and $v = 160\;{}^2\;{}^2 u$, we get from Equation (24) that with probability at-least $1 - \;=(2t^2(t+1))$,

$$\begin{aligned}
\sum_{s=1}^{-1} Y_s^{(u)} &\quad 8\;{}^2\ln\frac{2t^2(t+1)}{} 1_{>0} + \sqrt{\left(8\;{}^2\ln\frac{2t^2(t+1)}{} 1_{>0}\right)^2 + 160\;{}^2\;{}^2 u\ln\frac{2t^2(t+1)}{}};\\
&\overset{(a)}{\quad} 32\;{}^2\ln\frac{2t^2(t+1)}{}(\;+1)\sqrt{u+1}:
\end{aligned}$$

Step(a) follows from the fact that $\quad 1$. Thus, we have with probability at-least $1 - \frac{}{2t^2(t+1)}$,

$$\begin{aligned}
4\,{}_u^2 \sum_{s=1}^{-1} \frac{(u-s+\;)(\|{}_{u-s+1}^{(v)}\|_2^2 - E_{Z_{u-s+1}}[\|{}_{u-s+1}^{(v)}\|_2^2|F_{u-s+1}])}{u+\;} &\quad 96\,{}_u^2\;{}^2\ln\frac{2t^2(t+1)}{}(\;+1)\sqrt{u+1};\\
&\quad \frac{96\;{}^2\ln\frac{2t^2(t+1)}{}(\;+1)\sqrt{u+1}}{m^2(u+\;)^2};\\
&\quad \frac{96\;{}^2\ln\frac{2t^2(t+1)}{}(\;+1)}{m^2(u+\;)\sqrt{u+1}}:
\end{aligned}$$

Now taking an union bound over all $u \in \{1; \ldots; t\}$ yields that with probability at-least $1 - \frac{}{2t(t+1)}$, for all time $u \in \{1; \ldots; t\}$,

$$4\,{}_u^2 \sum_{s=1}^{-1} \frac{(t-s+\;)(\|{}_{u-s+1}^{(v)}\|_2^2 - E_{Z_{u-s+1}}[\|{}_{u-s+1}^{(v)}\|_2^2|F_{u-s+1}])}{t+\;} \quad \frac{96\;{}^2\ln\frac{2t^2(t+1)}{}(\;+1)}{m(u+\;)\sqrt{u+1}}$$

Proof of Equation (20)

$$2 \,{}_u \sum_{s=1}^{1} \frac{(u \quad s + \quad) h_{u \quad s} ; \;{}_{u \quad s+1}^{(v)} i}{u + } \qquad \frac{2}{m(u + \quad)^2} \sum_{s=1}^{1} h_{u \quad s} \qquad ; \;{}_{u \quad s+1}^{(v)} i$$

Fix a u 2 f 1; ; tg and denote by $Y_s^{(u)}$:= (u s +) $h_{u \quad s}$; ${}_{u \quad s+1}^{(v)}$ i . Since ${}_{u \quad s}$ is measurable with respect to the sigma-algebra generated by $F_{u \quad s}$, the conditional expectation $E[Y_s^{(u)} j F_{u \quad s}] = 0$. Thus, $(Y_s^{(u)})_{s=1}^{u \quad 1}$ is a martingale difference sequence with respect to the ltration $(F_{u \quad s})_{s=1}^{u \quad 1}$. Furthermore, we have from Equation (10) that $jY_s^{(u)} j$ 2(u s +) ${}_{u \quad s}^{(t)}$ + $\frac{R_1}{(u+ \quad 1)}$ 2 (u +) ${}_t^{(t)}$ + 2 G . We can now bound the sum of conditional variances as

$$\sum_{s=1}^{1} \mathrm{Var}(Y_s^{(u)} j F_{u \quad s}) \quad \sum_{s=1}^{1} 4(u \quad s + \quad)^2 ({}_{u \quad s}^{(t)})^2 \;\; ^2 \;\; ^2 + 4 \;\; ^2 G^2;$$

$$\overset{(18)}{} \; 12 \;\; ^2 \;\; ^2 (u + \quad)^2 (u + 1) \log(u + 1) ({}_u^{(t)})^2 + 4 \;\; ^2 \;\; ^2 G^2 u:$$

Step (a) follows since m < 1. Now applying the bound in Equation (24) with B := 2 (u +) ${}_u^{(t)}$ + 2 G and v = 12 ^{2} ^{2}(u +)^2(u + 1) log(u + 1)(${}_u^{(t)}$)^2 + 4 ^{2} ^{2}G^{2}u, we get that with probability at-least 1 =(2t^2(t + 1)) ,

$$\sum_{s=1}^{1} (u \quad s + \quad) h_{u \quad s} ; \;{}_{u \quad s+1}^{(v)} i \quad 2 \quad (u + \quad) \;{}_u^{(t)} + R_1 \;\; \ln \frac{2t^2(t+1)}{} \quad + \quad 2 \quad (u + \quad) \;{}_u^{(t)} + G \;\; \ln \frac{2t^2(t+1)}{} \;\; ^2$$

$$+ \;\; ^2 \;\; ^2 (u + \quad)^2 (u + 1) \log(u + 1) ({}_u^{(t)})^2 + 4 \;\; ^2 \;\; ^2 G^2 (u + 1) \quad \ln \frac{2t^2(t+1)}{} \;\; ^{\frac{1}{2}} ;$$

$$6(u + \quad) \sqrt{u+1} \log(u + 1) ({}_u^{(t)}) \quad (\quad + 1) \ln \frac{2t^2(t+1)}{} \quad + 2 \quad G \quad \sqrt{(u + 1) \ln \frac{2t^2(t+1)}{}} \; :$$

Thus,

$$2 \,{}_u \sum_{s=1}^{1} \frac{(u \quad s + \quad) h_{u \quad s} ; \;{}_{u \quad s+1}^{(v)} i}{u + } \qquad \frac{12 \sqrt{u+1} \log(u + 1) ({}_u^{(t)}) \quad (\quad + 1) \ln \frac{2t^2(t+1)}{}}{(u + \quad)} + \frac{C_t \; G}{10(u + 1)};$$

$$\frac{{}_u^{(t)} \ln \frac{2t^2(t+1)}{}}{10 \sqrt{u+1}} + \frac{C_t G}{10(u + 1)}:$$

The rst inequality follows since C_t $\frac{8 \sqrt{\ln \frac{2t^3}{}}}{^2 G}$. The last inequality follows since for all times u t, we have

$$\frac{12 \sqrt{u+1} \log(u + 1) \quad (\quad + 1) \ln \frac{2t^2(t+1)}{}}{(u + \quad)} \qquad \frac{\ln \frac{2t^2(t+1)}{}}{10}$$

as a consequence of 120 (+ 1) .

□

C PROOFS FROM SECTION 4.2

C.1 PROOF OF THEOREM 4.1

We bound this probability using the result of 3.1 and a simple union bound argument. For any Mr, observe that

$$P[9t \; 2 \; [r + 1 ; \;{}_c^{(r)}) \; s.t. A_t = 1 j A_r = 1] = P[[\,{}_{t=r+1}^{c \quad 1} A_t = 1 j A_r = 1]$$

$$\sum_{t=r+1}^{1} P[A_t = 1 | A_r = 1]: \tag{25}$$

We now examine the above Equation to bound it. For any fixed $t \in (r; \tau_c^{(r)})$

$$P[A_t = 1 | A_r = 1] = P\left[\bigcup_{s=r+1}^{t-1} \|\hat{b}_{r:s} - \hat{b}_{s+1:t}\| \geq B\left(s-r; \frac{\delta}{2t(t+1)}\right) + B\left(t-s-1; \frac{\delta}{2t(t+1)}\right)\right];$$

$$\leq \sum_{s=r+1}^{t-1} P\left(\|\hat{b}_{r:s} - \mu_{c-1}\| \geq B\left(s-r; \frac{\delta}{2t(t+1)}\right)\right) + P\left(\|\hat{b}_{s+1:t} - \mu_{c-1}\| \geq B\left(t-s-1; \frac{\delta}{2t(t+1)}\right)\right);$$

$$\stackrel{(a)}{\leq} \sum_{s=r+1}^{t-1} \frac{\delta}{2t(t+1)(s-r)(s-r+1)} + \frac{\delta}{2t(t+1)(t-s-1)(t-s)};$$

$$= \frac{\delta}{2t(t+1)}\left(\sum_{s=r+1}^{t-1} \frac{1}{(s-r)(s-r+1)} + \sum_{s=r+1}^{t-1} \frac{1}{(t-s-1)(t-s)}\right);$$

$$\leq \frac{\delta}{2t(t+1)}\left(\sum_{s=1}^{t-1-r} \frac{1}{s(s+1)} + \sum_{s=1}^{t-1-r} \frac{1}{s(s+1)}\right);$$

$$\stackrel{(b)}{\leq} \frac{\delta}{t(t+1)}: \tag{26}$$

Since for all $t < \tau_c^{(r)}$, the mean of the random variables $X_{r+1}; \cdots; X_t$ are identical and equal to μ_{c-1} (see notation in Section 2), Theorem 3.1 gives rise to inequality (a). Step (b) follows from the fact that $\sum_{s\geq 1} \frac{1}{s(s+1)} = 1$. Now substituting the bound from Equation (26) into Equation (25), we get that

$$P[\exists t \in [r+1; \tau_c^{(r)}) \text{ s.t. } A_t = 1 | A_r = 1] \leq \sum_{t=r+1}^{1} \frac{\delta}{t(t+1)};$$

$$\leq \sum_{t\geq 1} \frac{\delta}{t(t+1)};$$

$$= \delta:$$

Since the above bound holds for all r and process M, we have

$$\sup_{M;r} P[\exists t \in [r+1; \tau_c^{(r)}) \text{ s.t. } A_t = 1 | A_r = 1] \leq \delta:$$

C.2 PROOF OF LEMMA 4.2

Recall from the definition that the rth detection is false if

$$\rho_r^{(A)} = 1(\not\exists c \text{ s.t. } \tau_c \in (t_{r-1}^{(A)}; t_r^{(A)}]):$$

We will show that $E[\rho_r^{(A)}] \leq \delta$. This will then conclude the proof of the lemma.

$$E[\rho_r^{(A)}] = P[\not\exists c \text{ s.t. } \tau_c \in (t_{r-1}^{(A)}; t_r^{(A)}]];$$

$$= E\left[P[\not\exists c \text{ s.t. } \tau_c^{(s)} \in (s; t_r^{(A)}]] \,|\, t_{r-1}^{(A)} = s\right];$$

$$\leq E\left[P[\cup_{t=s+1}^{1} \tau_c^{(s)} = t; t_r^{(A)} < t] \,|\, t_{r-1}^{(A)} = s\right];$$

$$\mathbb{E}\quad \mathbb{P}[\exists t \in [s+1; \;\; c^{(s)}); A_t = 1] \quad t_{r\;\;1}^{(A)} = s \;\; ;$$

$$\overset{(a)}{}\quad \mathbb{E}\quad \mathbb{P}[\exists t \in [s+1; \;\; c^{(s)}); A_t = 1 | A_s = 1] \quad t_{r\;\;1}^{(A)} = s \;\; ;$$

$$\overset{(b)}{}\quad :$$

Inequality(a) follows from the fact that on the event $t_{r\;\;1}^{(A)} = s$, $A_s = 1$. Inequality(b) follows from Theorem 4.1.

D PROOF OF LEMMA 4.3

The proof follows from a straightforward application of Theorem 3.1 as follows. Let $n \in N$; $\;\; > 0$ and $\;^0 \in (0;1)$ be arbitrary.

$$\mathbb{P}[D(n; \;\; ; \;^0) \quad d] = \mathbb{P}[\cap_{s=1}^{n+d} A(X_{1:s}) = 0] ;$$

$$= \mathbb{P}\left[\bigcap_{s=1}^{n+d} \|\hat{b}_{1:s} \quad \hat{b}_{s+1:n+d}\|_2^2 \quad B \quad s; \frac{}{2(n+d)(n+d+1)} \quad + B \quad n+d \quad s \quad 1; \frac{}{2(n+d)(n+d+1)}\right] ;$$

$$\mathbb{P}\;\; \|\hat{b}_{1:n\;\;1} \quad \hat{b}_{n:n+d}\|_2^2 \quad B \quad n \quad 1; \frac{}{2(n+d)(n+d+1)} \quad + B \quad d; \frac{}{2(n+d)(n+d+1)} \quad : \tag{27}$$

From triangle-inequality, we know that

$$\|\hat{b}_{1:n\;\;1} \quad \hat{b}_{n:n+d}\|_2^2 \quad \| \;_1 \quad \;_2\|_2^2 \quad \|\hat{b}_{1:n\;\;1} \quad \;_1\|_2^2 \quad \|\hat{b}_{n:n+d} \quad \;_2\|_2^2;$$
$$= \;^2 \quad \|\hat{b}_{1:n\;\;1} \quad \;_1\|_2^2 \quad \|\hat{b}_{n:n+d} \quad \;_2\|_2^2: \tag{28}$$

Thus, substituting Equation (28 into Equation (27), we get that

$$\mathbb{P}[D(n; \;\; ; \;^0) \quad d] \quad \mathbb{P} \quad \;^2 \quad \|\hat{b}_{1:n\;\;1} \quad \;_1\|_2^2 \quad \|\hat{b}_{n:n+d} \quad \;_2\|_2^2$$
$$B \quad n \quad 1; \frac{}{2(n+d)(n+d+1)} \quad + B \quad d; \frac{}{2(n+d)(n+d+1)} \quad :$$

Denote by the events E_i for $i \in \{1;2\}$ as

$$E_1 := \quad \|\hat{b}_{1:n\;\;1} \quad \;_1\|_2^2 > B \quad n \quad 1; \frac{\;^0}{2} \quad ;$$

$$E_2 := \quad \|\hat{b}_{n:n+d} \quad \;_2\|_2^2 > B \quad d; \frac{\;^0}{2} \quad ;$$

Denote by $E := E_1 \cup E_2$. Theorem 3.1 gives that $\mathbb{P}[E_1] \quad \frac{\;^0}{2(n(n+1))} \quad \frac{\;^0}{2}$ and $\mathbb{P}[E_2] \quad \frac{\;^0}{2d(d+1)} \quad \frac{\;^0}{2}$. Thus, an union bound gives that $\mathbb{P}[E] \quad \;^0$. Let $d^0 \in G$ be arbitrary, where

$$G := \quad d \in N : \quad \;^2 \quad B \quad n \quad 1; \frac{\;^0}{2} \quad + B \quad d; \frac{\;^0}{2} \quad + B \quad n; \frac{}{2(n+d+1)(n+d)} \quad + B \quad d; \frac{}{2(n+d+1)(n+d)} \tag{29}$$

Claim : If the event E^c holds, then $D(n; \;\; ; \;\;) \quad d$ for all $d \in G$.

Suppose $d \in G$ and event E^c holds. Then, we know by triangle inequality in Equation (28) that

Figure 1: Plot of $D(n; \; ; \;^0)$ in Lemma 4.3 for xed $= 10 \; ; \; = 0{:}1$.

$$k\hat{b}_{1:n \; 1} \quad \hat{b}_{n:n+d}k_2^2 \quad k \quad _1 \quad _2k_2^2 \quad k \; \hat{b}_{1:n \; 1} \quad _1k_2^2 \quad k \; \hat{b}_{n:n+d} \quad _2k_2^2;$$

$$= \; ^2 \; k \; \hat{b}_{1:n \; 1} \quad _1k_2^2 \quad k \; \hat{b}_{n:n+d} \quad _2k_2^2; \tag{30}$$

$$\overset{(a)}{} \; ^2 \; B \quad n \quad 1; \frac{^0}{2} \quad B \quad d; \frac{^0}{2} \quad ; \tag{31}$$

$$\overset{(b)}{} B \quad n; \frac{}{2(n+d+1)(n+d)} + B \quad d; \frac{}{2(n+d+1)(n+d)} \quad : \tag{32}$$

Step (a) follows from the de nition of event E and on the assumption of the claim that event E^c holds. Step (b) follows from the fact that $d \; 2 \; G$ is arbitrary (cf. Equation (29). The last step says from Line 8 of Algorithm 1 that if no detection has been made till time $n+d$, then under the event E^c, time step d is a detection time. Since event E^c holds with probability at-least $1 \quad ^0$, this concludes the proof.

D.1 USEFUL CONVEXITY BASED INEQUALITIES

Let $f \; : \quad ! \quad R$ be a strongly convex function with strong convexity parameters $0 < m \quad M < 1$. Denote by $:= \arg\min_{\;2} \; f(\;)$. Since $f(\;)$ is convex and is convex and compact, the existence and uniqueness of is guaranteed. Strong convexity gives that for any $\hat{b}_{t \; 1} \; 2$,

$$f(\;) \quad f(\hat{b}_{t \; 1}) + hr\, f(\hat{b}_{t \; 1}); \quad \hat{b}_{t \; 1}i + \frac{m}{2}k \quad \hat{b}_{t \; 1}k_2^2: \tag{33}$$

Further since $= \arg\min_{\;2} \; f(\;)$., we have that

$$f(\hat{b}_{t \; 1}) \quad f(\;_t) \quad \frac{m}{2}k\hat{b}_{t \; 1} \quad k_2^2:$$

Putting these two together, we see that

$$hr\, f(\hat{b}_{t \; 1}); \hat{b}_{t \; 1} \quad i \quad mk\hat{b}_{t \; 1} \quad k_2^2: \tag{34}$$

Also, We further use the following lemma.

Lemma D.1 (Lemma 3:11 from [Bubeck, 2015].) Let $g : R^d \; ! \; R$ be a M smooth and m strongly convex function. Then for all $x; y \; 2 \; R^d$,

$$hr\, g(x) \quad r \quad g(y); x \quad yi \quad \frac{mM}{M+m}kx \quad yk_2^2 + \frac{1}{M+m}kr\, g(x) \quad r \quad g(y)k_2^2:$$

By substituting $x = \hat{b}_{t \; 1}$, $y = \;_t$ and $g(\;) = f(\;)$ and by leveraging the fact that $f(\;) = 0$, we get the inequality that

$$hr\, f(\hat{b}_{t \; 1}); \hat{b}_{t \; 1} \quad i \quad \frac{mM}{m+M}k\hat{b}_{t \; 1} \quad k_2^2 + \frac{1}{M+m}kr\, f(\hat{b}_{t \; 1})k_2^2:$$

Re-arranging, we see that

$$\|\nabla f(\widehat{\theta}_{t-1})\|_2^2 \leq (M+m)\langle \nabla f(\widehat{\theta}_{t-1}), \widehat{\theta}_{t-1} - \theta^* \rangle - mM\|\widehat{\theta}_{t-1} - \theta^*\|_2^2. \tag{35}$$

E ADDITIONAL SIMULATIONS

In Figure 2, we plot a sample path of observed data and mark out the true change-points and the detected time-instants by Algorithm 1. The plots indicate that although visually identifying the change in the means is hard, our change-point detection algorithm is able to consistently across variety of distribution families.

(a) Unit-variance Gaussian.

(b) Pareto with s = 2:1.

(c) Pareto with s = 2:01.

(d) Alternate Pareto s = 2:01 and Gaussian.

(e) Alternate Pareto s = 2:01 and Gaussian

(f) Alternate Pareto s = 2:01 and Gaussian

(g) Pareto s = 2:01; d = 15; = 5

(h) Pareto s = 2:01; d = 15; = 2

Figure 2: In all plots, we choose the change-point gap to be $\Delta = 0.1$ and $\delta = 0.05$ except (g) and (h) where $\Delta = 5$ and 2 respectively. In plots (g) and (h), we plot the norm of the observed random vector and thus the Y-axis is non-negative. We see missed detection in Figures (e) and (h) with the last change-point on the right being missed. We do not observe False-positives in these plots.