Supplementary Material

1 Lemmas for Proof of Theorem 4.1

Throughout this and the following section, we will assume the data is generated from the model detailed in Section 4.

We define part of the ratio in (4.2) as $X_k(\mu)$

$$X_k(\mu) = \exp\left\{\mu \sum_{u=t}^{s} \left(Y_{k,u} - \frac{\mu}{2}\right)\right\}.$$

The random variable $X_k(\mu)$ is log-normally distributed with different parameters depending on whether the sequence is normal or abnormal for that segment. In the normal segment case it is log-normal with parameters $-\mu^2(s-t+1)/2$ and $\mu^2(s-t+1)$,

$$\text{with } \mathbb{E}X_k(\mu) = 1.$$

For this case we will further define the mth central moment of $X_k(\mu)$ to be $C_m(\mu)$

$$C_m(\mu) = \mathbb{E}\left[(X_k(\mu) - \mathbb{E}X_k(\mu))^m\right].$$

Finally we denote the log of the product over the d terms in 4.2 as $S_d(\mu)$, taking the logarithm makes this become a sum over all the time-series

$$S_d(\mu) = \sum_{k=1}^{d} \log(1 + p_d(X_k(\mu) - 1)).$$

We now go on to prove several lemmas about $S_d(\mu)$ for both normal segments which will aid us in proving Theorems 4.1.

Lemma 1.1 (Normal segment moment bounds). *Assume we have a normal segment then*

$$\begin{aligned} \mathbb{E}S_d(\mu) &\leq -\frac{1}{2}C_2(\mu)dp_d^2 + \frac{1}{3}C_3(\mu)dp_d^3 \\ \mathbb{E}\left[(S_d(\mu) - \mathbb{E}S_d(\mu))^{2k}\right] &\leq K_k(\mu)d^k p_d^{2k} \end{aligned} \tag{1.1}$$

where $C_m(\mu)$ is the mth central moment of $X_m(\mu)$, and $K_k(\mu) > 0$ does not depend on d.

Proof. Writing out the expectation of $S_d(\mu)$ gives

$$\mathbb{E}S_d(\mu) = \sum_{k=1}^{d} \mathbb{E}\left[\log(1 + p_d(X_k(\mu) - 1))\right] \tag{1.2}$$

then we use the inequality $\log(1+x) \le x - \frac{x^2}{2} + \frac{x^3}{3}$ for $x > 0$. So

$$\begin{aligned}
\sum_{k=1}^{d} \mathbb{E}\left[\log(1+p_d(X_k(\mu)-1))\right] &\le \sum_{k=1}^{d} \mathbb{E}\left[p_d(X_k(\mu)-1) - \frac{p_d^2(X_k(\mu)-1)^2}{2} + \frac{p_d^3(X_k(\mu)-1)^3}{3}\right] \\
&= \sum_{k=1}^{d} -p_d^2 \frac{\mathbb{E}\left[(X_k(\mu)-1)^2\right]}{2} + p_d^3 \frac{\mathbb{E}\left[(X_k(\mu)-1)^3\right]}{3} \\
&= -\frac{1}{2} C_2(\mu) d p_d^2 + \frac{1}{3} C_3(\mu) d p_d^3.
\end{aligned}$$

Now to derive the second inequality we consider $S_d(\mu) - \mathbb{E}S_d(\mu)$

$$S_d(\mu) - \mathbb{E}S_d(\mu) = \sum_{i=1}^{d} \left[Z_i(\mu) - \mathbb{E}Z_i(\mu)\right] = \sum_{i=1}^{d} \bar{Z}_i(\mu),$$

where $\bar{Z}_i(\mu) = Z_i(\mu) - \mathbb{E}Z_i(\mu)$. Writing this in terms of the centered random variables $\bar{Z}_i(\mu)$ is advantageous as when we consider raising the sum to the $2k$th power any term including a unit power of $\bar{Z}_i(\mu)$ vanishes by independence as $\mathbb{E}\bar{Z}_i(\mu) = 0$. Define

$$\mathcal{I}_{d,k} = \left\{ (j_1, \ldots, j_d) : j_i \in \{0, 2, 3, \ldots, 2k\} \text{ for } i = 1, \ldots, d \text{ and } \sum_{i=1}^{d} j_i = 2k \right\},$$

the set of non-negative integer vectors of length d, whose entries sum to $2k$, and that have no-entry that is equal to 1. For $\mathbf{j} \in \mathcal{I}_{d,k}$, let $n_{\mathbf{j}}$ be the number of terms in the expansion of $(\sum_{i=1}^{d} \bar{Z}_i(\mu))^{2k}$ which have powers j_i for $\bar{Z}_i(\mu)$). Thus

$$\begin{aligned}
\mathbb{E}\left[(S_d(\mu) - \mathbb{E}S_d(\mu))^{2k}\right] &= \mathbb{E}\left[\left(\sum_{i=1}^{d} \bar{Z}_i(\mu)\right)^{2k}\right] \\
&= \sum_{\mathbf{j} \in \mathcal{I}_{d,k}} n_{\mathbf{j}} \prod_{i=1}^{d} \mathbb{E}\left(\bar{Z}_i(\mu)^{j_i}\right) \\
&\le \mathbb{E}\left(\bar{Z}_1(\mu)^{2k}\right) \sum_{\mathbf{j} \in \mathcal{I}_{d,k}} n_{\mathbf{j}}.
\end{aligned}$$

Using $|\log(1+x)| \le |x| + x^2/2$, we can bound $\mathbb{E}(\bar{Z}_1^{2k})$ by $A_k(\mu) p_d^{2k}$, where $A_k(\mu)$ will depend only on the the first $2k$ moments of $X_k(\mu)$, but not on p_d. Finally note that each term in $\mathcal{I}_{d,k}$ can only involve vectors with at most k non-zero components. For a term with l non-zero-components there will be $O(d^l)$ possible choices for which components are non-zero. Hence we have that

$$\sum_{\mathbf{j} \in \mathcal{I}_{d,k}} n_{\mathbf{j}} \le B_k d^k,$$

for some constant B_k that does not depend on d. Thus we have the required result, with $K_k(\mu) = A_k(\mu) B_k$. $\square$

Lemma 1.2 (Probability bound)**.** *Fix μ and assume $p_d \to 0$ as $d \to \infty$. For a normal segment we have that there exists $D_k(\mu) > 0$ such that for sufficiently large d*

$$\Pr\left(S_d(\mu) \geq -\frac{1}{4}C_2(\mu)dp_d^2\right) \leq \frac{D_k(\mu)}{d^k p_d^{2k}}. \tag{1.3}$$

Proof. We first bound the probability by the absolute value of the centered random variable and then use Markov's inequality with an even power of the form $2k$

$$\begin{aligned}\Pr\left(S_d(\mu) \geq -\frac{1}{4}C_2(\mu)dp_d^2\right) &\leq \Pr\left(|S_d(\mu) - \mathbb{E}S_d(\mu)| \geq \frac{1}{4}C_2(\mu)dp_d^2 - \frac{1}{3}C_3(\mu)dp_d^3\right)\\ &\leq \frac{\mathbb{E}\left[(S_d(\mu) - \mathbb{E}S_d(\mu))^{2k}\right]}{(\frac{1}{4}C_2(\mu)dp_d^2 - \frac{1}{3}C_3(\mu)dp_d^3)^{2k}}.\end{aligned}$$

For d sufficiently large that $2C_3(\mu)p_d < C_2(\mu)$, we have

$$\frac{1}{4}C_2(\mu)dp_d^2 - \frac{1}{3}C_3(\mu)dp_d^3 > \frac{1}{12}C_2(\mu)dp_d^2.$$

Now using the result from Lemma 1.1 we can replace the $2k$th centered moment by the bound we obtained above. Thus for sufficiently large d,

$$\Pr\left(S_d(\mu) \geq -\frac{1}{4}C_2(\mu)dp_d^2\right) \leq \frac{K_k(\mu)d^k p_d^{2k}}{(\frac{1}{12}C_2(\mu)dp_d^2)^{2k}}$$

So the result holds with $D_k(\mu) = K_k(\mu)[C_2(\mu)/12]^{-2k}$. □

Lemma 1.3 (Lower bound for the second derivative of $S_d(\mu)$)**.** *We have that*

$$\frac{d^2 S_d(\mu)}{d\mu^2} \geq -d(s-t+1)$$

Proof. Firstly note that

$$\frac{\mathrm{d}X_k(\mu)}{\mathrm{d}\mu} = \left(\sum_{u=t}^{s} y_{k,u} - \mu(s-t+1)\right) X_k(\mu).$$

Now differentiating $S_d(\mu)$ twice

$$\begin{aligned}\frac{\mathrm{d}S_d(\mu)}{\mathrm{d}\mu} &= \sum_{k=1}^{d} \frac{p_d\left(\sum_{u=t}^{s} y_{k,u} - \mu(s-t+1)\right) X_k(\mu)}{1 + p_d(X_k(\mu) - 1)}\\ \frac{\mathrm{d}^2 S_d(\mu)}{\mathrm{d}\mu^2} &= \sum_{k=1}^{d} \frac{-p_d(s-t+1)X_k(\mu) + p_d\left(\sum_{u=t}^{s} y_{k,u} - \mu(s-t+1)\right)^2 X_k(\mu)}{1 + p_d(X_k(\mu) - 1)}\\ &\quad - \left(\sum_{u=t}^{s} y_{k,u} - \mu(s-t+1)\right)^2 \left(\frac{p_d X_k(\mu)}{1 + p_d(X_k(\mu) - 1)}\right)^2\end{aligned}$$

Let

$$Q_k = \frac{p_d X_k(\mu)}{1 + p_d(X_k(\mu) - 1)}$$

and $0 \le Q_k \le 1$ as $1 - p_d > 0$ (or $p_d < 1$). Thus the second derivative

$$\begin{aligned}\frac{\mathrm{d}^2 S_d(\mu)}{\mathrm{d}\mu^2} &= \sum_{k=1}^{d} \left[-(s-t+1)Q_k + \left(\sum_{u=t}^{s} y_{k,u} - \mu(s-t+1) \right)^2 (Q_k - Q_k^2) \right] \\ &\ge \sum_{k=1}^{d} -(s-t+1)Q_k \ge -d(s-t+1)\end{aligned}$$

has the required lower bound. □

Lemma 1.4 (Detection of normal segments)**.** *Let $\pi(\mu)$ be a density function with support $[a,b]$ with $a > 0$ and $b < \infty$, and assume $1/p_d = O(d^{\frac{1}{2}-\epsilon})$ for some $\epsilon > 0$. For a normal segment $[t,s]$,*

$$\int \left\{ \prod_{k=1}^{d} \frac{P_{A,k}(t,s;\mu)}{P_{N,k}(t,s)} \right\} \pi(\mu) d\mu \to 0 \tag{1.4}$$

in probability as $d \to \infty$.

Proof. Define $C_2 = \min_{\mu \in [a,b]} C_2(\mu)$, and for a given d, M_d to be the smallest integer that is greater than

$$\frac{(b-a)\sqrt{s-t+1}}{p_d\sqrt{C_2}}.$$

Define $\Delta_d = (b-a)/M_d$. Now we can partition $[a,b]$ into M_d intervals of the form $[\mu_{i-1}, \mu_i]$ for $i = 1, \dots, M_d$, where $\mu_i = a + i\Delta_d$. Then the left-hand side of (1.4) can be rewritten as

$$\sum_{i=1}^{M_d} \int_{\mu_{i-1}}^{\mu_i} \left\{ \prod_{k=1}^{d} [1 + p_d(X_k(\mu) - 1)] \right\} \pi(\mu) \mathrm{d}\mu.$$

Remember that $S_d(\mu) = \sum_{k=1}^{d} \log[1 + p_d(X_k(\mu) - 1)]$. Let E_d be the event that

$$S_d(\mu) < -\frac{1}{4} C_2 d p_d^2, \text{for all } \mu = \mu_i,\ i = 0, \dots, M_d.$$

If this event occurs then

$$\max_{\mu \in [a,b]} S_d(\mu) < -\frac{1}{4} C_2 d p_d^2 + \Delta_d^2 d(s-t+1)/8,$$

as using Lemma 1.3 we can bound $S_d(\mu)$ on each interval $[\mu_i, \mu_{i+1}]$ by a quadratic with second derivative $-d(s-t+1)$ and which takes values $-\frac{1}{4} C_2 d p_d^2$ at the end-points.

Now by definition of Δ_d,

$$-\frac{1}{4}C_2 dp_d^2 + \Delta_d^2 d(s-t+1)/8 < -\frac{1}{4}C_2 dp_d^2 + \frac{1}{8}C_2 dp_d^2 \to -\infty$$

as $d \to \infty$ because $dp_d^2 \to \infty$ under our assumption on p_d. Thus to prove the Lemma we need only show that event E_d occurs with probability 1 as $d \to \infty$.

We can bound the probability of E_d not occurring using Lemma 1.2. For any integer $k > 0$ we have that the probability E_d does not occur is

$$\begin{aligned}
\sum_{i=1}^{M_d+1} \Pr\left(S_d(\mu_i) \geq -\frac{1}{4}C_2 dp_d^2\right) &\leq \sum_{i=1}^{M_d+1} \Pr\left(S_d(\mu_i) \geq -\frac{1}{4}C_2(\mu_i) dp_d^2\right) \\
&\leq \sum_{i=1}^{M_d+1} \frac{D_k(\mu_i)}{d^k p_d^{2k}} \\
&\leq (M_d+1) \max_{\mu\in[a,b]} \frac{D_k(\mu)}{d^k p_d^{2k}}.
\end{aligned}$$

Here $D_k(\mu)$ is defined in Lemma 1.2. It is finite for any μ, and hence $\max_{\mu\in[a,b]} D_k(\mu)$ is finite.

Now $M_d = O(p_d^{-1})$, so we have that the above probability is $O(d^{-k} p_d^{-2k-1}) = O(d^{1/2-(2k+1)\epsilon})$. So by choosing $k > 1/(4\epsilon)$ this is $O(d^{-\epsilon})$ which tends to 0 as required. □

2 Lemmas for Proof of Theorem 4.2

We use the same notation as in Section 4.1. However, we will now consider an abnormal segment from positions t to s. Let α_d denote the proportion of sequences that are abnormal, and μ_0 the mean. The observations in this segment come from a two component mixture. With probability α_d they are normally distributed with mean μ_0 and variance 1; otherwise they have a standard normal distribution. It is straightforward to show that for such an abnormal segment,

$$\mathbb{E}X_k(\mu) = (1-\alpha_d) + \alpha_d e^{\mu\mu_0(s-t+1)}. \tag{2.1}$$

Lemma 2.1 (Abnormal segments, expectation and variance)**.** *Assume we have an abnormal segment* $[t,s]$ *with the mean of affected dimensions being* μ_0*. Let* $f(\mu)$ *be a density function with support* $A \subset \mathbb{R}$ *then*

$$\mathbb{E}\left[\int_A S_d(\mu) f(\mu)\, d\mu\right] \geq D_1(\mu) dp_d$$

$$\mathrm{Var}\left(\int_A S_d(\mu) f(\mu)\, d\mu\right) \leq D_2(\mu) dp_d^2 + o(dp_d^2)$$

with

$$D_1(\mu) \quad = \quad \min_{\mu\in A}\left(\mathbb{E}[X_k(\mu)-1] - \frac{p_d}{2}\mathbb{E}[(X_k(\mu)-1)^2]\right) \tag{2.2}$$

$$
\begin{aligned}
&= \min_{\mu\in A}\left[\alpha_d(e^{\mu\mu_0(s-t+1)}-1)-\frac{p_d}{2}\left(e^{\mu^2(s-t+1)}-1\right)-\frac{\alpha_d p_d C(\mu)}{2}\right] \quad (2.3)\\
C(\mu) &= e^{\mu^2(s-t+1)}(e^{2\mu\mu_0(s-t+1)}-1)-2(e^{\mu\mu_0(s-t+1)}-1)
\end{aligned}
$$

and

$$D_2(\mu)=\max_{\mu\in A}\mathbb{E}\left[(X_k(\mu)-1)^2\right].$$

Proof. As $S_d(\mu)$ is the sum of d iid terms we can rewrite the expectation and variance with a single term

$$
\begin{aligned}
\mathbb{E}\left[\int_A S_d(\mu)f(\mu)\mathrm{d}\mu\right] &= d\mathbb{E}\left[\int_A \log(1+p_d(X_k(\mu)-1))f(\mu)\mathrm{d}\mu\right]\\
\mathrm{Var}\left(\int_A S_d(\mu)f(\mu)\mathrm{d}\mu\right) &= d\mathrm{Var}\left(\int_A \log(1+p_d(X_k(\mu)-1))f(\mu)\mathrm{d}\mu\right).
\end{aligned}
$$

Now as $\log(1+x)>x-x^2/2$,

$$
\begin{aligned}
\mathbb{E}\left[\int_A \log(1+p_d(X_k(\mu)-1))f(\mu)\mathrm{d}\mu\right] &\geq \mathbb{E}\left[\int_A\left(p_d(X_k(\mu)-1)-\frac{p_d^2(X_k(\mu)-1)^2}{2}\right)f(\mu)\mathrm{d}\mu\right]\\
&= p_d\int_A\left(\mathbb{E}[X_k(\mu)-1]-\frac{p_d}{2}\mathbb{E}[(X_k(\mu)-1)^2]\right)f(\mu)\mathrm{d}\mu,
\end{aligned}
$$

which gives (2.2). We then obtain (2.3) by using (2.1) and a similar calculation for the variance of $X_k(\mu)$.

We now consider the variance, which is bounded by the second moment. Using $|\log(1+x)|\leq|x|+x^2/2$ we have

$$
\begin{aligned}
\mathrm{Var}\left(\int_A \log(1+p_d(X_k(\mu)-1))f(\mu)\mathrm{d}\mu\right) &\leq \mathbb{E}\left[\left(\int_A \log(1+p_d(X_k(\mu)-1))f(\mu)\mathrm{d}\mu\right)^2\right]\\
&\leq \mathbb{E}\left[\int_A\left\{\log(1+p_d(X_k(\mu)-1))\right\}^2 f(\mu)\mathrm{d}\mu\right]\\
&\leq \mathbb{E}\left[\int_A\left\{p_d^2(X_k(\mu)-1)^2+p_d^3|X_k(\mu)-1|^3+\frac{p_d^4}{4}(X_k(\mu)-1)^4\right\}f(\mu)\mathrm{d}\mu\right]\\
&\leq \max_{\mu\in A}\mathbb{E}\left\{p_d^2(X_k(\mu)-1)^2\right\}\int_A f(\mu)\mathrm{d}\mu+o(p_d^2),
\end{aligned}
$$

which gives the required bound for the variance. ☐

Lemma 2.2 (Detection of abnormal segments)**.** *Assume that we have an abnormal segment* $[t,s]$*. Let* α_d *be the probability of a sequence being abnormal and the mean of the abnormal observations be* μ_0*, with* $p_d=o(1)$*. Assume that there exists a set* A *such that for all* $\mu\in A$ *we have*

$$\lim_{d\to\infty}\alpha_d\left(e^{\mu\mu_0(s-t+1)}-1\right)-\frac{p_d}{2}\left(e^{\mu^2(s-t+1)}-1\right)>\delta,$$

and $\int_A \pi(\mu)d\mu > \delta'$, *for some* $\delta, \delta' > 0$. *If* $dp_d^2 \to \infty$ *as* $d \to \infty$ *then*

$$\int \left\{ \prod_{k=1}^{d} \frac{P_{A,k}(t,s;\mu)}{P_{N,k}(t,s)} \right\} \pi(\mu)\, d\mu \to \infty \tag{2.4}$$

in probability as $d \to \infty$.

Proof. If we restrict the integral in (2.4) to one over $A \subset \mathbb{R}$ we get a lower bound. Then rewriting the ratio in (2.4), using (4.2), in terms of $X_k(\mu)$ we get

$$\int \left\{ \prod_{k=1}^{d} [1 + p_d(X_k(\mu) - 1)] \right\} \pi(\mu)\mathrm{d}\mu \geq \int_A \left\{ \prod_{k=1}^{d} [1 + p_d(X_k(\mu) - 1)] \right\} \pi(\mu)\mathrm{d}\mu.$$

If we consider the logarithm of the above random variable and use Jensen's inequality we get a lower bound

$$\begin{aligned} \log \left(\int_A \left\{ \prod_{k=1}^{d} [1 + p_d(X_k(\mu) - 1)] \right\} \pi(\mu)\mathrm{d}\mu \right) &\geq \int_A \left\{ \sum_{k=1}^{d} \log(1 + p_d(X_k(\mu) - 1)) \right\} \pi(\mu)\mathrm{d}\mu \\ &= \int_A S_d(\mu)\pi(\mu)\mathrm{d}\mu. \end{aligned}$$

Then if we can show this random variable goes to ∞ as $d \to \infty$ the original random variable has the same limit. Let $T_d = \int_A S_d(\mu)\pi(\mu)\mathrm{d}\mu$. Using Lemma 2.1, we have

$$\mathrm{E}(T_d) > \log(\delta') + \delta d p_d,$$

and for sufficiently large d there exists a constant C such that

$$\mathrm{Var}(T_d) < C d p_d^2.$$

So by Chebyshev's inequality

$$\Pr(T_d \leq \log(\delta') + \delta d p_d - d p_d^2) \leq \Pr(|T_d - \mathbb{E}T_d| \geq d p_d^2) \leq \frac{\mathrm{Var}(T_d)}{d^2 p_d^4} < \frac{C}{d p_d^2}.$$

Thus $T_d \to \infty$ in probability as $d \to \infty$, which implies (2.4). $\square$

Supplementary Material (Supplement A).