Supplementary material for Transformed-Linear Models for Time Series Extremes

Nehali Mhatre
Department of Statistics
Colorado State University
nmhatre@colostate.edu

Daniel Cooley
Department of Statistics
Colorado State University
cooleyd@stat.colostate.edu

Summary

This supplementary material shows that the space $\mathbb{V}$ of infinite transformed-linear combinations of regularly-varying random variables with absolutely summable coefficients is a vector space, it shows the conditions of inner product are met, and shows $\mathbb{V}$ is isomorphic to the vector space of absolutely summable sequences ℓ^1. This supplementary material also demonstrates through simulation, the bias reduction in TPDF estimation by subtracting off the mean of the marginally transformed Fréchet time series.

1 Proof of Inner Product space

1.1 Vector Space $\mathbb{V}$

Consider the space $\mathbb{V} = \{X_t : X_t = \bigoplus_{j=0}^{\infty} \psi_{t,j} \circ Z_j, \sum_{j=0}^{\infty} |\psi_j| < \infty\}$ where Z_j's are independent and tail stationary $RV_+(2)$ random variables with $\Pr(Z_j > x)/(x^{-2}L(x)) = 1$ and $\psi_{t,j} \in \mathbb{R}$. We first show that $\mathbb{V}$ is a vector space. Consider X_t, X_u, $X_v \in \mathbb{V}$, $a, b \in \mathbb{R}$ and $\mathbb{O} := \bigoplus_{j=0}^{\infty} 0 \circ Z_j$. Let $f : \mathbb{R} \to (0, \infty)$ defined as $f(y) = \log\{1 + \exp(y)\}$ be the transform defined in Cooley and Thibaud (2019).

<u>Vector addition is closed:</u> $X_t \oplus X_u \in \mathbb{V}$.

Proof:

$$
\begin{aligned}
X_t \oplus X_u &= \bigoplus_{j=0}^{\infty} \psi_{t,j} \circ Z_j \oplus \bigoplus_{j=0}^{\infty} \psi_{u,j} \circ Z_j = f\left\{\sum_{j=0}^{\infty} \psi_{t,j} f^{-1}(Z_j)\right\} \oplus f\left\{\sum_{j=0}^{\infty} \psi_{u,j} f^{-1}(Z_j)\right\} \\
&= f\left(f^{-1}\left[f\left\{\sum_{j=0}^{\infty} \psi_{t,j} f^{-1}(Z_j)\right\}\right] + f^{-1}\left[f\left\{\sum_{j=0}^{\infty} \psi_{u,j} f^{-1}(Z_j)\right\}\right]\right) \\
&= f\left\{\sum_{j=0}^{\infty} \psi_{t,j} f^{-1}(Z_j) + \sum_{j=0}^{\infty} \psi_{u,j} f^{-1}(Z_j)\right\}, \qquad (1) \\
&= f\left\{\sum_{j=0}^{\infty} (\psi_{t,j} + \psi_{u,j}) f^{-1}(Z_j)\right\} = \bigoplus_{j=0}^{\infty} (\psi_{t,j} + \psi_{u,j}) \circ (Z_j) \qquad (2)
\end{aligned}
$$

which is in $\mathbb{V}$, since $\sum_{j=0}^{\infty} |\psi_{t,j} + \psi_{u,j}| \leq \sum_{j=0}^{\infty} |\psi_{t,j}| + \sum_{j=0}^{\infty} |\psi_{u,j}| < \infty$.

<u>Vector addition is commutative:</u> $X_t \oplus X_u = X_u \oplus X_t$.

Proof: By (1),

$$
\begin{aligned}
X_t \oplus X_u &= f\left\{\sum_{j=0}^{\infty}\psi_{t,j}f^{-1}(Z_j)+\sum_{j=0}^{\infty}\psi_{u,j}f^{-1}(Z_j)\right\}\\
&= f\left\{\sum_{j=0}^{\infty}\psi_{u,j}f^{-1}(Z_j)+\sum_{j=0}^{\infty}\psi_{t,j}f^{-1}(Z_j)\right\} = X_u \oplus X_t.
\end{aligned}
$$

<u>Vector addition is associative:</u> $X_t \oplus (X_u \oplus X_v) = (X_t \oplus X_u) \oplus X_v$.

Proof: By (2),

$$
\begin{aligned}
X_t \oplus (X_u \oplus X_v) &= f\left\{\sum_{j=0}^{\infty}\psi_{t,j}f^{-1}(Z_j)\right\} \oplus f\left\{\sum_{j=0}^{\infty}\psi_{u,j}f^{-1}(Z_j)+\sum_{j=0}^{\infty}\psi_{v,j}f^{-1}(Z_j)\right\}\\
&= f\left(f^{-1}\left[f\left\{\sum_{j=0}^{\infty}\psi_{t,j}f^{-1}(Z_j)\right\}\right]+f^{-1}\left[f\left\{\sum_{j=0}^{\infty}\psi_{u,j}f^{-1}(Z_j)+\sum_{j=0}^{\infty}\psi_{v,j}f^{-1}(Z_j)\right\}\right]\right)\\
&= f\left\{\sum_{j=0}^{\infty}\psi_{t,j}f^{-1}(Z_j)+\sum_{j=0}^{\infty}\psi_{u,j}f^{-1}(Z_j)+\sum_{j=0}^{\infty}\psi_{v,j}f^{-1}(Z_j)\right\}\\
&= f\left(f^{-1}\left[f\left\{\sum_{j=0}^{\infty}\psi_{t,j}f^{-1}(Z_j)+\sum_{j=0}^{\infty}\psi_{u,j}f^{-1}(Z_j)\right\}\right]+f^{-1}\left[f\left\{\sum_{j=0}^{\infty}\psi_{v,j}f^{-1}(Z_j)\right\}\right]\right)\\
&= f\left\{\sum_{j=0}^{\infty}\psi_{t,j}f^{-1}(Z_j)+\sum_{j=0}^{\infty}\psi_{u,j}f^{-1}(Z_j)\right\} \oplus f\left\{\sum_{j=0}^{\infty}\psi_{v,j}f^{-1}(Z_j)\right\}\\
&= (X_t \oplus X_u) \oplus X_v.
\end{aligned}
$$

<u>Additive identity:</u> $X_t \oplus \mathbb{O} = X_t$.

Proof: By (2),

$$X_t \oplus \mathbb{O} = \bigoplus_{j=0}^{\infty}\psi_{t,j}\circ Z_j \oplus \bigoplus_{j=0}^{\infty} 0\circ Z_j = \bigoplus_{j=0}^{\infty}(\psi_{t,j}+0)\circ Z_j = \bigoplus_{j=0}^{\infty}\psi_{t,j}\circ Z_j = X_t.$$

<u>Additive inverse:</u> $X_t \oplus -X_t = \mathbb{O}$.

Proof:

$$X_t \oplus -X_t = \bigoplus_{j=0}^{\infty}\psi_{t,j}\circ Z_j \oplus \bigoplus_{j=0}^{\infty} -\psi_{t,j}\circ Z_j = \bigoplus_{j=0}^{\infty}(\psi_{t,j}+(-\psi_{t,j}))\circ Z_j = \bigoplus_{j=0}^{\infty} 0\circ Z_j = \mathbb{O}.$$

<u>Scalar multiplication is closed:</u> $a \circ X_t \in \mathbb{V}$.

Proof:

$$
\begin{aligned}
a\circ X_t = a\circ\bigoplus_{j=0}^{\infty}\psi_{t,j}\circ Z_j &= f\left(a\cdot f^{-1}\bigoplus_{j=0}^{\infty}\psi_{t,j}\circ Z_j\right)\\
&= f\left(a\cdot f^{-1}\left[f\left\{\sum_{j=0}^{\infty}\psi_{t,j}f^{-1}(Z_j)\right\}\right]\right) = f\left\{a\sum_{j=0}^{\infty}\psi_{t,j}f^{-1}(Z_j)\right\}\\
&= f\left\{\sum_{j=0}^{\infty}a\psi_{t,j}f^{-1}(Z_j)\right\} = \bigoplus_{j=0}^{\infty}(a\psi_{t,j})\circ Z_j \in \mathbb{V}. \qquad (3)
\end{aligned}
$$

Scalar multiplication is distributive: $a \circ (X_t \oplus X_u) = a \circ X_t \oplus a \circ X_u$.

Proof: By (1),

$$\begin{aligned}
a \circ (X_t \oplus X_u) &= a \circ f\left\{\sum_{j=0}^{\infty} \psi_{t,j} f^{-1}(Z_j) + \sum_{j=0}^{\infty} \psi_{u,j} f^{-1}(Z_j)\right\} \\
&= f\left(af^{-1}\left[f\left\{\sum_{j=0}^{\infty} \psi_{t,j} f^{-1}(Z_j) + \sum_{j=0}^{\infty} \psi_{u,j} f^{-1}(Z_j)\right\}\right]\right) \\
&= f\left\{a\sum_{j=0}^{\infty} \psi_{t,j} f^{-1}(Z_j) + a\sum_{j=0}^{\infty} \psi_{u,j} f^{-1}(Z_j)\right\} \\
&= f\left(af^{-1}\left[f\left\{\sum_{j=0}^{\infty} \psi_{t,j} f^{-1}(Z_j)\right\}\right] + af^{-1}\left[f\left\{\sum_{j=0}^{\infty} \psi_{u,j} f^{-1}(Z_j)\right\}\right]\right) \\
&= f\left\{af^{-1}\left(\bigoplus_{j=0}^{\infty} \psi_{t,j} \circ Z_j\right) + af^{-1}\left(\bigoplus_{j=0}^{\infty} \psi_{u,j} \circ Z_j\right)\right\} \\
&= a \circ X_t \oplus a \circ X_u.
\end{aligned}$$

Scalar multiplication is associative: $a \circ (b \circ X_t) = (ab) \circ X_t$.

Proof: By (3)

$$\begin{aligned}
a \circ (b \circ X_t) = a \circ \left\{\bigoplus_{j=0}^{\infty} (b\psi_{t,j}) \circ Z_j\right\} &= \left\{\bigoplus_{j=0}^{\infty} (ab\psi_{t,j}) \circ Z_j\right\} \\
&= (ab) \circ \left(\bigoplus_{j=0}^{\infty} \psi_{t,j} \circ Z_j\right) = (ab) \circ X_t. \qquad (4)
\end{aligned}$$

Multiplicative identity: $1 \circ X_t = X_t$.

Proof: By (3),

$$1 \circ X_t = \bigoplus_{j=0}^{\infty} (1 \cdot \psi_{t,j}) \circ Z_j = \bigoplus_{j=0}^{\infty} \psi_{t,j} \circ Z_j = X_t.$$

1.2 Inner Product in $\mathbb{V}$

Let X_t and X_s be two elements of vector space $\mathbb{V}$ such that $X_t = \bigoplus_{j=0}^{\infty} \psi_{t,j} \circ Z_j$ and $X_s = \bigoplus_{j=0}^{\infty} \psi_{s,j} \circ Z_j$ where $\sum_{j=0}^{\infty} |\psi_{t,j}| < \infty$ and $\sum_{j=0}^{\infty} |\psi_{s,j}| < \infty$. We define the inner product between X_t and X_s as,

$$\langle X_t, X_s \rangle := \sum_{j=0}^{\infty} \psi_{t,j} \psi_{s,j}. \qquad (5)$$

We can show that (5) is indeed an inner product by showing that it satisfies the properties of an inner product. Consider $X_t, X_u, X_v \in \mathbb{V}$ and $a \in \mathbb{R}$.

Linearity: $\langle X_t \oplus X_u, X_v \rangle = \langle X_t, X_v \rangle + \langle X_u, X_v \rangle$ and $\langle a \circ X_t, X_v \rangle = a\langle X_t, X_v \rangle$.

Proof: By (2),

$$\begin{aligned}
\langle X_t \oplus X_u, X_v \rangle &= \left\langle \bigoplus_{j=0}^{\infty} (\psi_{t,j} + \psi_{u,j}) \circ Z_j,\ \bigoplus_{j=0}^{\infty} \psi_{v,j} \circ Z_j \right\rangle = \sum_{j=0}^{\infty} (\psi_{t,j} + \psi_{u,j})\psi_{v,j} \\
&= \sum_{j=0}^{\infty} \psi_{t,j}\psi_{v,j} + \sum_{j=0}^{\infty} \psi_{u,j}\psi_{v,j} = \langle X_t, X_v \rangle + \langle X_u, X_v \rangle.
\end{aligned}$$

Also,

$$\langle a \circ X_t, X_v \rangle = \left\langle a \circ \bigoplus_{j=0}^{\infty} \psi_{t,j} \circ Z_j,\ \bigoplus_{j=0}^{\infty} \psi_{v,j} \circ Z_j \right\rangle = \left\langle \bigoplus_{j=0}^{\infty} (a\psi_{t,j}) \circ Z_j,\ \bigoplus_{j=0}^{\infty} \psi_{v,j} \circ Z_j \right\rangle$$
$$= \sum_{j=0}^{\infty} a\psi_{t,j}\psi_{v,j} = a\sum_{j=0}^{\infty} \psi_{t,j}\psi_{v,j} = a\langle X_t, X_v \rangle$$

Symmetric property: $\langle X_t, X_u \rangle = \langle X_u, X_t \rangle$.
Proof:

$$\langle X_t, X_u \rangle = \sum_{j=0}^{\infty} \psi_{t,j}\psi_{u,j} = \sum_{j=0}^{\infty} \psi_{u,j}\psi_{t,j} = \langle X_u, X_t \rangle.$$

Positive definite property: $\langle X_t, X_t \rangle \geq 0$ and $\langle X_t, X_t \rangle = 0$ if and only if $X_t = \mathbb{O} := \bigoplus_{j=0}^{\infty} 0 \circ Z_j$.
Proof:

$$\langle X_t, X_t \rangle = \sum_{j=0}^{\infty} \psi_{t,j}\psi_{t,j} = \sum_{j=0}^{\infty} \psi_{t,j}^2 \geq 0.$$

Let $\langle X_t, X_t \rangle = 0$, which means that $\sum_{j=0}^{\infty} \psi_{t,j}^2 = 0$. For $\psi_{t,j} \in \mathbb{R}$, this is possible only if $\psi_{t,j} = 0$ for all j, equivalently, when $X_t = \mathbb{O}$. Conversely, if $X_t = \mathbb{O}$, $\langle X_t, X_t \rangle = \sum_{j=0}^{\infty} 0^2 = 0$.

Thus, by satisfying all the above properties, the vector space $\mathbb{V}$ is an inner product space. For $X_t \in \mathbb{V}$, the norm of X_t is defined as $\|X_t\| = \sqrt{\langle X_t, X_t \rangle} = \sqrt{\sum_{j=0}^{\infty} \psi_{t,j}^2}$, which is finite as $\sum_{j=0}^{\infty} |\psi_{t,j}| < \infty$. Also, $X_t, X_s \in \mathbb{V}$ are said to be orthogonal if $\langle X_t, X_s \rangle = \sum_{j=0}^{\infty} \psi_{t,j}\psi_{s,j} = 0$.

As vector space $\mathbb{V}$ is an inner product space, it follows that $\mathbb{V}$ is a metric space. We can define a metric given by the squared norm of $X_t \ominus X_s$ as,

$$\langle X_t \ominus X_s,\ X_t \ominus X_s \rangle = \sum_{j=0}^{\infty} (\psi_{t,j} - \psi_{s,j})^2.$$

1.3 Isomorphism of $\mathbb{V}$ to ℓ^1

Consider the infinite dimensional space of absolutely summable sequences,

$$\ell^1 = \left\{ \{a_j\}_{j=0}^{\infty}, a_j \in \mathbb{R} : \sum_{j=0}^{\infty} |a_j| < \infty \right\}.$$

Vector addition and scalar multiplication in ℓ^1 are defined component-wise: $(a_0, a_1, a_2, ...) + (b_0, b_1, b_2, ...) = (a_0 + b_0, a_1 + b_1, a_2 + b_2, ...)$ for $\{a_j\}, \{b_j\} \in \ell^1$ and the product of $(a_0, a_1, a_2, ...)$ with a scalar $\alpha \in \mathbb{R}$ is the sequence $(\alpha a_0, \alpha a_1, \alpha a_2, ...)$.

For any $X_t \in \mathbb{V}$ we can define a mapping $T : \mathbb{V} \to \ell^1$ such that $T(X_t) = \{\psi_{t,j}\}_{j=0}^{\infty} \in \ell^1$. We will first show that the mapping T is a linear map. Let X_t and X_s be two elements of vector space $\mathbb{V}$ such that $X_t = \bigoplus_{j=0}^{\infty} \psi_{t,j} \circ Z_j$ and $X_s = \bigoplus_{j=0}^{\infty} \psi_{s,j} \circ Z_j$ where $\sum_{j=0}^{\infty} |\psi_{t,j}| < \infty$ and $\sum_{j=0}^{\infty} |\psi_{s,j}| < \infty$. By (2) and (3), for any scalars $a, b \in \mathbb{R}$,

$$\begin{aligned} T(a \circ X_t \oplus b \circ X_s) &= T\left\{ \bigoplus_{j=0}^{\infty} (a\psi_{t,j}) \circ Z_j \oplus \bigoplus_{j=0}^{\infty} (b\psi_{s,j}) \circ Z_j \right\} = T\left\{ \bigoplus_{j=0}^{\infty} (a\psi_{t,j} + b\psi_{s,j}) \circ Z_j \right\} \\ &= \{a\psi_{t,j} + b\psi_{s,j}\}_{j=0}^{\infty} = \{a\psi_{t,j}\}_{j=0}^{\infty} + \{b\psi_{s,j}\}_{j=0}^{\infty} \\ &= a\{\psi_{t,j}\}_{j=0}^{\infty} + b\{\psi_{s,j}\}_{j=0}^{\infty} = aT(X_t) + bT(X_s). \end{aligned} \tag{6}$$

Recall that a linear map T is called an isomorphism if T is one-to-one and onto. We will show that the linear map T defined above is an isomorphism. Let X_t and X_s be as defined above. Then,

$$T(X_t) = T(X_s) \implies \{\psi_{t,j}\}_{j=0}^{\infty} = \{\psi_{s,j}\}_{j=0}^{\infty} \implies t = s \implies X_t = X_s. \tag{7}$$

Thus T is a one-to-one map. T is also onto since, for any sequence $\{a_j\}_{j=0}^{\infty}$ in ℓ^1 we can define $X = \bigoplus_{j=0}^{\infty} a_j \circ Z_j$ which is in $\mathbb{V}$ as $\sum_{j=0}^{\infty} |a_j| < \infty$. Thus, $\mathbb{V}$ is isomorphic to ℓ^1.

2 Bias in TPDF Estimation of Transformed Regularly-Varying Time Series

2.1 Background

Let $\{X_t\}$ be a regularly varying tail stationary time series with $\alpha = 2$. Let $(X_t, X_{t+h})^T$ be a two-dimensional random vector of elements at lag h of $\{X_t\}$. Thus, there exists a function $b(s)$ and Radon measure $H_{X_t,X_{t+h}}$ on $\Theta_1^+ = \{\mathbf{w} = (w_t, w_{t+h}) \in \bar{\mathbb{X}}^2 : \|\mathbf{w}\|_2 = 1\}$ such that as $s \to \infty$,

$$s\text{Pr}\left(\frac{(X_t, X_{t+h})}{b(s)} \in \cdot\right) \xrightarrow{v} \nu_{X_t,X_{t+h}}(\cdot), \text{ where } \nu_{X_t,X_{t+h}}(dr \times d\mathbf{w}) = 2r^{-3} dr dH_{X_t,X_{t+h}}(\mathbf{w}). \tag{8}$$

We also assume that the lower tail condition

$$s\text{Pr}\{X_i \leq \exp(-kb(s))\} \to 0, \quad k > 0, \quad i = 1, ..., p, \quad s \to \infty, \tag{9}$$

is met. We define the tail pairwise dependence function (TPDF) as

$$\sigma(X_t, X_{t+h}) = \int_{\Theta_1^+} w_t w_{t+h} dH_{X_t,X_{t+h}}(\mathbf{w}).$$

To estimate the tail pairwise dependence function (TPDF), we use the estimator defined in Cooley and Thibaud (2019) in which the true angular measure is replaced by an empirical estimate. Let $\{x_t\}$, $(t = 1, \cdots, n)$, be the time series observations. Let $(x_t, x_{t+h})^T$, $(t = 1, \cdots, n-h)$ be lag-h pairs of observations from $\{x_t\}$. Let $r_t = \|(x_t, x_{t+h})^T\|_2$, and $\mathbf{w}_t = (w_t, w_{t+h})^T = (x_t, x_{t+h})^T / r_t$. The TPDF estimator is defined as

$$\hat{\sigma}(h) = \hat{m} \int_{\Theta_1^+} w_t w_{t+h} d\hat{N}_{X_t,X_{t+h}}(w) = \frac{\hat{m}}{\sum_{t=1}^{n} \mathbb{I}(r_t > r_0)} \sum_{t=1}^{n} w_t w_{t+h} \mathbb{I}(r_t > r_0), \tag{10}$$

where r_0 is some high threshold for the radial components, $N_{X_t,X_{t+h}}(\cdot) = m^{-1} H_{X_t,X_{t+h}}(\cdot)$, and $\hat{m}$ is an estimate of $H_{X_t,X_{t+h}}(\Theta_1^+)$. Because we preprocessed the time series to have a unit scale, $m = 2$ and does not need to be estimated. When the data are not preprocessed to have a unit scale, an empirical estimator is $\hat{m} = \frac{r_0^2}{n} \sum_{t=1}^{n} \mathbb{I}(r_t > r_0)$.

It is known that tail dependence estimates tend to have positive bias in the case of weak dependence (Huser et al., 2016). In this supplementary material, we explore the bias associated with estimation of the TPDF. Simulation study shows that subtracting off the mean of the series, considerably reduces bias in TPDF estimation.

2.2 Simulated data

First we consider a transformed regularly-varying MA(2) time series

$$X_t = Z_t \oplus \theta_1 \circ Z_{t-1} \oplus \theta_2 \circ Z_{t-2},$$

where $\{Z_t\}$ is a sequence of independent positive Fréchet random numbers with unit scale and shape $\alpha = 2$, for four different pairs of values of θ_1 and θ_2. Similarly, we simulate a transformed regularly-varying MA(5) time series

$$X_t = Z_t \oplus \theta_1 \circ Z_{t-1} \oplus \cdots \oplus \theta_5 \circ Z_{t-5},$$

with two different sets of values for $\theta_1, \cdots, \theta_5$, and a transformed regularly-varying AR(1) time series

$$X_t = \phi \circ X_{t-1} \oplus Z_t,$$

with four different values of ϕ. We simulate 50,000 observations x_t, of each of the above time series. For each of the time series, we then obtain the transformed observations $x_t^{(\text{trans})} = g(x_t)$, where g is the inverse cdf of the desired Fréchet. That is, $g(x) = \{-\log\hat{F}(x)\}^{-1/2}$, so that $\text{Pr}(X_t^{(\text{trans})} \leq x) \approx \exp(-x^{-2})$, and $X_t^{(\text{trans})} \in RV_+(2)$ with scale 1.

2.3 TPDF estimation

We estimate the TPDF for the first 10 lags for the MA time series, for the first 50 lags for the AR time series with positive ϕ, and for the first 20 lags for the AR time series with negative ϕ, using the estimator in (10). TPDF estimation is done in four different ways. Method 1: $\hat{m}$ in (10) is estimated as $\frac{r_0^2}{n}\sum_{t=1}^{n}\mathbb{I}(r_t > r_0)$. Method 2: m is considered to be 2, as it theoretically should be after preprocessing the time series to have a unit scale. Method 3: Mean of the time series $X_t^{\text{(trans)}}$ is subtracted off before estimation and m is estimated as in Method 1. Method 4: Mean of the time series $X_t^{\text{(trans)}}$ is subtracted off before estimation and m is considered to be 2. In method 3 and 4, after we subtract off the mean of the time series, we set the negative values to zero.

The theoretical TPDF for the unit scale MA time series is given by $\frac{\sum_{j=0}^{q}\theta_j^{(0)}\theta_{j+h}^{(0)}}{\sum_{j=0}^{q}\theta_j^{(0)2}}$, where $q = 2$ or 5 for the MA(2) and MA(5) time series respectively, and that for the unit scale AR(1) time series is $\left(\phi^h\right)^{(0)}$, where $a^{(0)} = \max(a, 0)$.

2.4 Results

Figure 1 shows the theoretical and estimated TPDF using the four methods described in Section 2.3 for different transformed MA time series. It can be seen that there is substantial positive bias in TPDF estimation by method 1 and 2. When we subtract off the mean of the time series in method 3 and 4, the bias reduces considerably. However, there is still some bias at higher lags where the theoretical TPDF is zero. Also, estimating m in method 3 underestimates the TPDF substantially. We see similar results in Figure 2 for the transformed AR time series.

When we simulate data using any regularly-varying noise terms, there always exists a bias problem. Our aim is to minimize the bias problem. A lag-1 bivariate scatter plot of the time series observations will have a bunch of large points away from the axes. As the lag increases, the points on the scatter plot spread out and move towards the axes. If there is no long range dependence, those points should eventually align along the axes, but that does not happen. However, the closer we can make the irrelevant points to zero, the better is the behavior of our estimator. Subtracting off the mean of the time series, reduces the bias in TPDF estimation. This can be seen in Figure 3 which gives lag-5 plots for the MA(2) time series without subtracting the mean off (left panel) and after subtracting the mean off (right panel). We can clearly identify the few points towards the middle of the graph in both the plots, that have a higher value for both of the axes. Subtracting off the mean does not affect these points in the middle of the graph. However, it drives the points near the axes, closer to the axes, thus reducing bias in estimation.

2.5 Conclusion

Simulation results indicate that subtracting off the mean of the marginally transformed Fréchet time series considerably reduces bias in TPDF estimation. Also, considering $m = 2$ instead of estimating m gives better TPDF estimates. There is still bias in the tail which can be further reduced by subtracting the bias off of the TPDF at higher lags. But it is tricky to address the bias this way since selecting the lag, beyond which to subtract off the bias, will be subjective. However, after subtracting off the mean of the time series, the TPDF estimates at lower lags where dependence exists, are not too different from their theoretical counterparts.

References

Cooley, D. and Thibaud, E. (2019). Decompositions of dependence for high-dimensional extremes. *Biometrika.* **106**(3):587–604.

Huser, R., Davison, A. C. & Genton, M. G. (2016). Likelihood estimators for multivariate extremes. *Extremes.* **19**(1), 79–103.

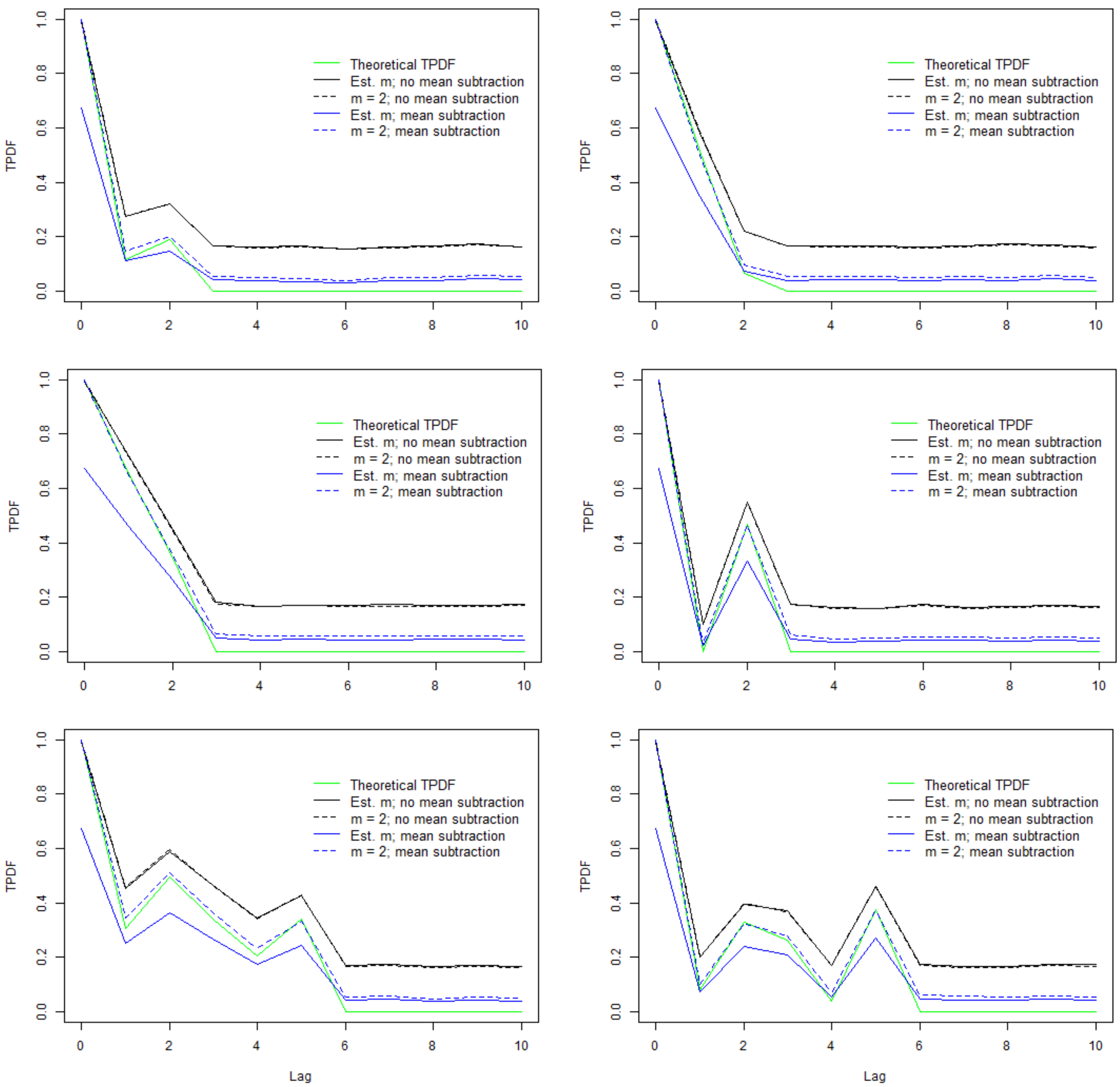

Figure 1: Comparison of theoretical TPDF and estimated TPDF for transformed regularly-varying MA time series. MA(2): $\theta_1 = 0.1, \theta_2 = 0.2$ (upper left panel), MA(2): $\theta_1 = 0.7, \theta_2 = 0.1$ (upper right panel), MA(2): $\theta_1 = 0.9, \theta_2 = 0.9$ (middle left panel), MA(2): $\theta_1 = -0.5, \theta_2 = 0.7$ (middle right panel), MA(5): $\theta_1 = 0.1, \theta_2 = 0.7, \theta_3 = 0.2, \theta_4 = 0.4, \theta_5 = 0.8$ (lower left panel), and MA(5): $\theta_1 = 0.1, \theta_2 = 0.7, \theta_3 = -0.2, \theta_4 = -0.4, \theta_5 = 0.8$ (lower right panel).

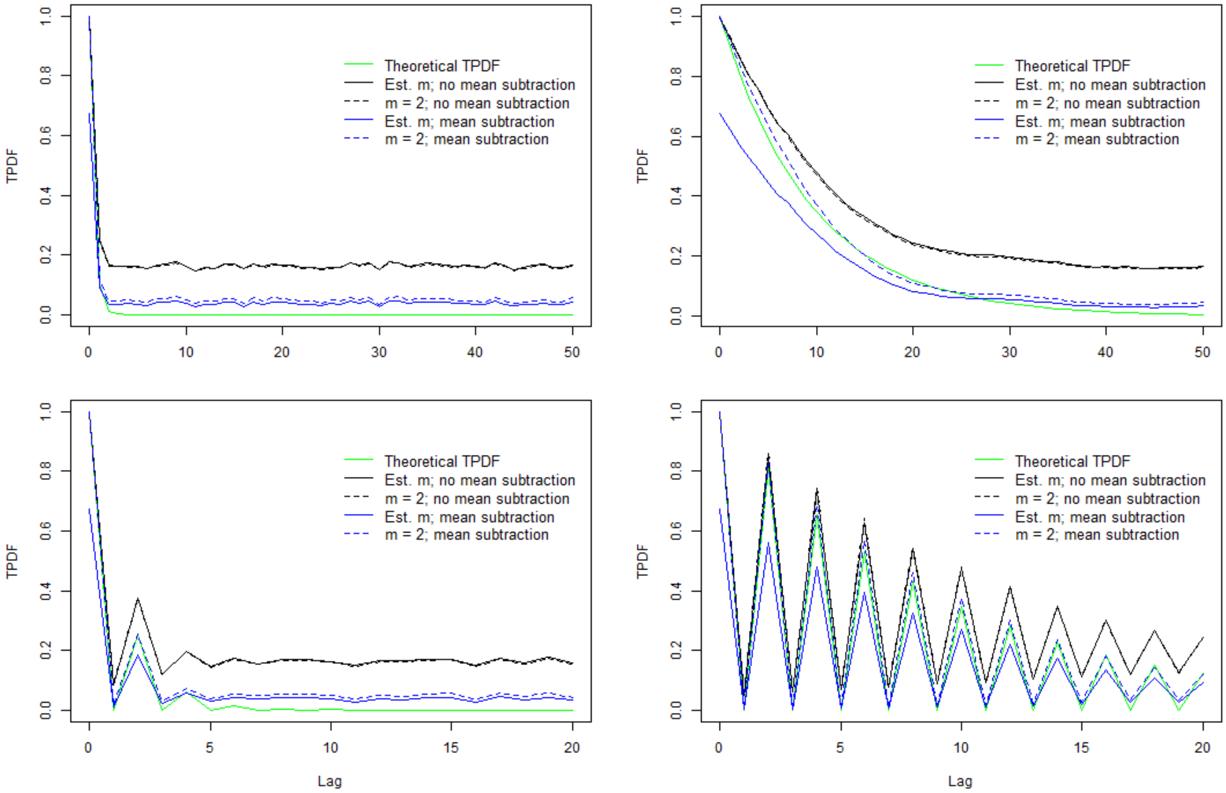

Figure 2: Comparison of theoretical TPDF and estimated TPDF for transformed regularly-varying AR(1) time series. $\phi = 0.1$ (upper left panel), $\phi = 0.9$ (upper right panel), $\phi = -0.5$ (lower left panel), and $\phi = -0.9$ (lower right panel).

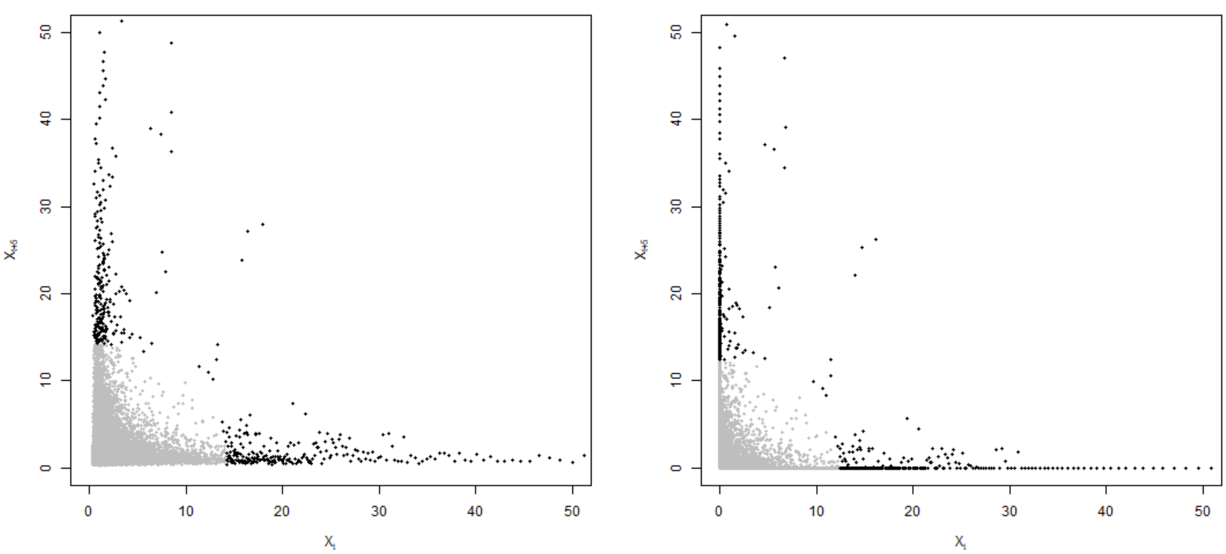

Figure 3: Lag-5 plots of the MA(2) time series with $\theta_1 = 0.9$ and $\theta_2 = 0.9$. No mean subtracted from the time series (left panel), mean subtracted from the time series (right panel).