SUPPLEMENTARY MATERIAL

Online Supplementary:

(doi: COMPLETED BY THE TYPESETTER; .pdf). The online supplementary material contains the detailed proofs of Theorem 2.1 and some useful lemmas. The long detailed steps are in section 6 and the lemmas are postponed to section 7.

6. Detailed Steps of the Proof of Theorem 2.1

6.1. Preparation stage: The preparation stage consists of truncation approximation, m-dependence approximation and blocking approximation.

6.1.1. Truncation approximation: Truncation approximation is necessary to allow higher moments manipulations. For $b > 0$ and $v = (v_1, \ldots, v_d)^T \in \mathbb{R}^d$, define

$$T_b(v) = (T_b(v_1), \ldots, T_b(v_d))^T,$$

where $T_b(w) = \min(\max(w, -b), b)$. (6.1)

**Proposition 6.1.** Assume Condition (2.A). It is possible to choose a sequence $t_n \to 0$ slow enough such that we have

$$\max_{1 \leq i \leq n} |S_i - S_i^{\oplus}| = o_P(n^{1/p}), \text{ where } S_i^{\oplus} = \sum_{i=1}^{t} [T_{t_n^{1/p}}(X_i) - ET_{t_n^{1/p}}(X_i)].$$

(6.2)

**Proof.** of Proposition 6.1. We introduce a very slowly converging sequence $t_n \to 0$ based on the uniform integrability condition (2.A). For every $t > 0$, we have

$$\sup_i \frac{1}{t^p} E(|X_i|^p 1_{|X_i| > t n^{1/p}}) = 0 \text{ and } n \sup_i E \min\left(\frac{|X_i|^\gamma}{t^\gamma n^{\gamma/p}}, 1\right) \to 0 \text{ as } n \to \infty,$$

(6.3)
where $\gamma > p$. The second relation follows from Lemma 7.1. Clearly (6.3) implies that
\begin{equation}
\sup_i \frac{1}{t_n} E(|X_i|^p 1_{|X_i| > t_n n^{1/p}}) + n \sup_i E \min(\frac{|X_i|^\gamma}{t_n n^{\gamma/p}}, 1) \to 0 \text{ as } n \to \infty, \tag{6.4}
\end{equation}
holds for a sequence $t_n \to 0$ very slowly. Without loss of generality we can let
\begin{equation}
t_n \log \log n \to \infty \tag{6.5}
\end{equation}
since otherwise we can replace $t_n$ by $\max(t_n, (\log \log n)^{-1/2})$ (say). The truncation operator $T_b$ in (6.1) is Lipschitz continuous with Lipschitz constant 1. Let
\begin{equation}
R_{c,l} = \sum_{i=1}^{l+c} X_i^+ = \sum_{i=1}^{l+c} [T_{t_n n^{1/p}}(X_i) - E(T_{t_n n^{1/p}}(X_i))]. \tag{6.6}
\end{equation}

By (6.4), we have $P(\max_{i \leq n} |S_i - \sum_{j=1}^i T_{t_n n^{1/p}}(X_j)| = 0) \to 1$ in view of
\begin{equation*}
\sup_j P(|X_j| > t_n n^{1/p}) \leq \sup_j \frac{1}{nt_n} E \left( |X_j|^p I \left( |X_j| > t_n n^{1/p} \right) \right) = o(1/n).
\end{equation*}
Also by (6.4), $\max_{j \leq n} |E(X_j - T_{t_n n^{1/p}}(X_j))| = o(n^{1/p-1})$. Hence (6.2) follows. \hfill \Box

6.1.2. $m$-dependence approximation: The $m$-dependence approximation is a very important tool that is extensively used in literature; see for example the Gaussian approximation in Liu and Lin (2009, [13]) and Berkes, Liu and Wu (2014, [2]). For a suitably chosen sequence $m$, we look at the conditional mean $E(X_i | \epsilon_{i}, \ldots, \epsilon_{i-m})$. This gives a very simple yet effective way to handle the original process in terms of a collection of $\epsilon_i$'s. Define the partial sum process
\begin{equation}
\tilde{R}_{c,l} = \sum_{i=1}^{l+c} \tilde{X}_j, \text{ where } \tilde{X}_j = E(T_{t_n n^{1/p}}(X_j) | \epsilon_j, \ldots, \epsilon_{j-m}) - E(T_{t_n n^{1/p}}(X_j)). \tag{6.7}
\end{equation}
Write $\tilde{R}_{0,i} = \tilde{S}_i$. From Lemma A1 in Liu and Lin (2009, [13]), we have
\[
\| \max_{1 \leq i \leq n} |S_i^\oplus - \tilde{S}_i|_r \| \leq c_r n^{1/2} \Theta_{1+m,r}.
\] (6.8)

The proofs in [13] are for stationary processes. Since our $\delta_{j,r}$ in (2.1) is defined in an uniform manner, the proof goes through for the non-stationary case as well. Assume
\[
n^{1/2 - 1/r} \Theta_{m,r} \to 0.
\] (6.9)

By (6.8) and (6.9), we have $n^{1/r}$ convergence in the $m$-dependence approximation step
\[
\max_{1 \leq i \leq n} |S_i^\oplus - \tilde{S}_i| = o_P(n^{1/r}).
\] (6.10)

6.1.3. Blocking approximation: Towards the blocking approximation, we approximate the partial sum process $\tilde{S}_i$ by sums of $A_j$ where, for $j \geq 0$,
\[
A_{j+1} = \sum_{i=k_0 m+1}^{(2k_0 j + 2k_0)m} \tilde{X}_i, \text{ where } k_0 = \left\lfloor \Theta_{0,2}/\lambda_* \right\rfloor + 2.
\] (6.11)

To this end, we will need the following two conditions, for some $\gamma > p$,
\[
n^{1-\gamma/r} m^{\gamma/2 - 1} \to 0,
\] (6.12)
\[
n^{1/p - 1/\gamma} \sum_{j=m+1}^{\infty} \delta_{j,p}^{\gamma} \to 0.
\] (6.13)
We now define functional dependence measure for the truncated process \((T_{t_n^{1/p}}(X_i))_{i \leq n}\) as

\[
\delta_{j,l}^\oplus = \sup_i \|T_{t_n^{1/p}}(X_i) - T_{t_n^{1/p}}(X_{i,(i-j)})\|_l, \text{ where } l \geq 2.
\]

Similarly, define the functional dependence measure for the \(m\)-dependent process \((\tilde{X}_i)\) as

\[
\tilde{\delta}_{j,l} = \sup_i \|\tilde{X}_i - \tilde{X}_{i,(i-j)}\|_l.
\]

For these dependence measures, the following inequality holds for all \(l \geq 2\):

\[
\tilde{\delta}_{j,l} \leq \delta_{j,l}^\oplus \leq \delta_{j,l}.
\] (6.14)

We now proceed to proving Proposition 6.2, the blocking approximation result. As mentioned in the main text, we need to assume conditions (6.12) and (6.13) for this step. The almost-polynomial rate of \(m\) sequence as mentioned in (6.15) is also assumed.

Remark: We need another condition for the blocking approximation (see (7.2) in the proof of Lemma 7.3). However, we skip it here and choose \(m\) and \(\gamma\) such that conditions (6.9), (6.12) and (6.13) are met. These will automatically imply this fourth one in view of (2.3).

We assume an almost polynomial rate for \(m\) sequence: for some \(0 < L < 1\),

\[
m = \lfloor n^{L t_n^k} \rfloor, \quad 0 < k < (\gamma - p)/(\gamma/2 - 1).
\] (6.15)
Proposition 6.2. Assume (6.12) and (6.13) for some $\gamma > p$. Moreover, assume (6.15) for the $m$ sequence and (2.3) for the decay rate of $\Theta_{i,p}$ with some $A > \gamma/p$.

Then

$$\max_{1 \leq i \leq n} |S_i - S_i^\circ| = o_P(n^{1/r}), \text{ where } S_i^\circ = \sum_{j=1}^{q_i} A_j, \quad q_i = \lceil i/(2k_0m) \rceil.$$  \hspace{1cm} (6.16)

Proof. of Proposition 6.2: Let $S = \{2ik_0m, 0 \leq i \leq q_n\}$, $\phi_n = (n^{1-\gamma/r}m^{\gamma/2-1})^{1/(2r)}$.

Then

$$P \left( \max_{1 \leq i \leq n} |\tilde{R}_{0,l} - \sum_{j=1}^{\lceil i/(2k_0m) \rceil} A_j| \geq \phi_n n^{1/r} \right) \leq \frac{n}{2k_0m} \max_{c \in S} P(\max_{1 \leq i \leq 2k_0m} |\tilde{R}_{c,l}| \geq \phi_n n^{1/r})$$

$$\leq n \max_{c \in S} \frac{E(\max_{1 \leq i \leq 2k_0m} |\tilde{R}_{c,l}|)}{2k_0m\phi_n^{1/r}} = O(\phi_n^r),$$

from the assumption (6.12) and Lemma 7.3. Since $\phi_n \to 0$, (6.16) follows. \hfill \Box

Summarizing (6.2), (6.10) and (6.16), we can work on $S_i^\circ$ in view of

$$\max_{1 \leq i \leq n} |S_i - S_i^\circ| = o_P(n^{1/r}).$$  \hspace{1cm} (6.17)

In the next two subsections we shall provide details of the arguments for steps mentioned in sections 4.2 and 4.3. Section 6.2 presents the conditional Gaussian approximation, where we shall apply Proposition 6.3 stated in section 7. Section 6.3 deals with unconditional Gaussian approximation and regrouping.
6.2. Conditional Gaussian approximation: The blocks $A_j$ created in (6.11) after the blocking approximation are weakly independent; except they share some dependence on the border. In this subsection, we look at the conditional process given the $\epsilon_i$ the blocks share in their borders. Demeaning the conditional process, we apply the Proposition 6.3 for the Gaussian approximation. For $1 \leq i \leq n$, let $\tilde{H}_i$ be a measurable function such that

$$\tilde{X}_i = \tilde{H}_i(\epsilon_i, \ldots, \epsilon_{i-m}). \quad (6.18)$$

Recall Proposition 6.2 for the definition of $q_i$. Let $q = q_n$. For $j = 1, \ldots, q$, define

$$\bar{a}_{2k_0j} = \{a_{(2k_0j-1)m+1}, \ldots, a_{2k_0jm}\} \text{ and } a = \{\ldots, \bar{a}_{2k_0}, \bar{a}_{4k_0}, \ldots\}.$$

Given $a$, define, for $2k_0jm + 1 \leq i \leq (2k_0j + 1)m$,

$$\tilde{X}_i(\bar{a}_{2k_0j}) = \tilde{H}_i(\epsilon_i, \ldots, \epsilon_{2k_0jm+1}, a_{2k_0jm}, \ldots, a_{i-m})$$

and for $(2k_0j + 2k_0 - 1)m + 1 \leq i \leq (2k_0j + 2k_0)m$,

$$\tilde{X}_i(\bar{a}_{2k_0j+2k_0}) = \tilde{H}_i(a_i, \ldots, a_{(2k_0j+2k_0-1)m+1}, \epsilon_{(2k_0j+2k_0-1)m}, \ldots, \epsilon_{i-m}).$$
Further, define the blocks as following,

\[ F_{4j+1}(\bar{a}_{2k_0j}) = \sum_{i=(2k_0jm+1)}^{(2k_0jm+1)m} \bar{X}_i(\bar{a}_{2k_0j}), \]  \hspace{1cm} (6.19) \]

\[ F_{4j+2} = \sum_{i=(2k_0jm+1)m+1}^{(2k_0jm+2k_0)m} \bar{X}_i, \quad F_{4j+3} = \sum_{i=(2k_0jm+k_0)m+1}^{(2k_0jm+k_0)m+1} \bar{X}_i, \]

\[ F_{4j+4}(\bar{a}_{2k_0j+2k_0}) = \sum_{i=(2k_0jm+k_0m+1)}^{(2k_0jm+k_0m+2k_0)m} \bar{X}_i(\bar{a}_{2k_0j+2k_0}). \]

Similarly, for \( j = 1, \ldots, q \), define

\[ \bar{\vartheta}_{2k_0j} = \{ \epsilon(2k_0j-1)m+1, \ldots, \epsilon(2k_0jm) \} \] and \( \vartheta = \{ \ldots, \bar{\vartheta}_0, \bar{\vartheta}_{2k_0}, \bar{\vartheta}_{4k_0}, \ldots \}. \)

Recall \( A_j \) from (6.11). We have

\[ A_{j+1} = F_{4j+1}(\bar{\vartheta}_{2k_0j}) + F_{4j+2} + F_{4j+3} + F_{4j+4}(\bar{\vartheta}_{2k_0j+2k_0}). \]

Define the mean functions

\[ \Lambda_{4j+1}(\bar{a}_{2k_0j}) = E^*(F_{4j+1}(\bar{a}_{2k_0j})) \] and \( \Lambda_{4j+4}(\bar{a}_{2k_0j+2k_0}) = E^*(F_{4j+4}(\bar{a}_{2k_0j+2k_0})) \),

where \( E^* \) refers to the conditional moment given \( a \). In the sequel, with slight abuse of notation, we will simply use the usual \( E \) to denote moments of random variables conditioned on \( a \). Introduce the centered process

\[ Y_j(\bar{a}_{2k_0j}, \bar{a}_{2k_0j+2k_0}) = F_{4j+1}(\bar{a}_{2k_0j}) - \Lambda_{4j+1}(\bar{a}_{2k_0j}) + F_{4j+2} + F_{4j+3} + F_{4j+4}(\bar{a}_{2k_0j+2k_0}) - \Lambda_{4j+4}(\bar{a}_{2k_0j+2k_0}). \]  \hspace{1cm} (6.20) \]
Following the definition of $S_n^\circ$, we let

$$S_i(a) = \sum_{j=0}^{q-1} Y_j(\bar{a}_{2k_0j}, \bar{a}_{2k_0j+2k_0}).$$

The mean and variance function of $S_i(a)$ are respectively denoted by

$$M_i(a) = \sum_{j=0}^{q-1} [\Lambda_{4j+1}(\bar{a}_{2k_0j}) + \Lambda_{4j+4}(\bar{a}_{2k_0j+2k_0})],$$

$$Q_i(a) = \sum_{j=0}^{q-1} V_j(\bar{a}_{2k_0j}, \bar{a}_{2k_0j+2k_0}),$$

where $V_j(\bar{a}_{2k_0j}, \bar{a}_{2k_0j+2k_0})$ is the dispersion matrix of $Y_j(\bar{a}_{2k_0j}, \bar{a}_{2k_0j+2k_0})$. Define

$$V_{j0}(\bar{a}_{2k_0j}) = E(F_{4j-2}F_{4j-1}^T + F_{4j-1}F_{4j-2}^T) + Var(F_{4j-1} + F_{4j}(\bar{a}_{2k_0j}) - \Lambda_{4j}(\bar{a}_{2k_0j})) + Var(F_{4j+1}(\bar{a}_{2k_0j}) - \Lambda_{4j+1}(\bar{a}_{2k_0j}) + F_{4j+2}).$$

Note that, the following identity holds for all $t$:

$$\sum_{j=0}^{t} V_j(\bar{a}_{2k_0j}, \bar{a}_{2k_0j+2k_0}) = L(\bar{a}_0) + \sum_{j=1}^{t-1} V_{j0}(\bar{a}_{2k_0j}) + U_t(\bar{a}_{2k_0j+2k_0}),$$

where $L(\bar{a}_0) = Var(F_1(\bar{a}_0) + F_2)$ and

$$U_{t-1}(\bar{a}_{2k_0t}) = E(F_{4t-2}F_{4t-1}^T + F_{4t-1}F_{4t-2}^T) + Var(F_{4t-1} + F_{4t}(\bar{a}_{2k_0t}) - \Lambda_{4t}(\bar{a}_{2k_0t})).$$

Define

$$L^a_{\gamma} = \sum_{j=0}^{q-1} E(|Y_j(\bar{a}_{2k_0j}, \bar{a}_{2k_0j+2k_0})|^\gamma).$$
In the sequel, we suppress $Y_j(\bar{a}_{2k_0j}, \bar{a}_{2k_0j} + 2k_0)$, $V_j(\bar{a}_{2k_0j}, \bar{a}_{2k_0j} + 2k_0)$, $Y_j(\bar{\vartheta}_{2k_0j}, \bar{\vartheta}_{2k_0j} + 2k_0)$, $V_j(\bar{\vartheta}_{2k_0j}, \bar{\vartheta}_{2k_0j} + 2k_0)$, $V_j(\bar{2k_0j}, \bar{2k_0j} + 2k_0)$ and $V_j(\bar{\vartheta}_{2k_0j})$ as just $Y^a_j, Y^\vartheta_j, V^a_j, V^\vartheta_j$ and $V^\vartheta_j$ respectively. We apply Proposition 6.3 to the independent mean zero random vectors $Y^a_j$.

Proposition 6.3 concerns Gaussian approximation for independent vectors. There are several types of Gaussian approximations in literature for independent vectors. We find the following result by Götze and Zaitsev (2008, [10]) particularly useful since it provides an explicit and good approximation bound for the partial sums. This has been used several times in our proof.

**Proposition 6.3.** Let $\xi_1, \ldots, \xi_n$ be independent $\mathbb{R}^d$-valued mean zero random vectors. Assume that there exist $s \in \mathbb{N}$ and a strictly increasing sequence of non-negative integers $\eta_0 = 0 < \eta_1 < \ldots < \eta_s = n$ satisfying the following conditions.

Let

$$\zeta_k = \xi_{\eta_{k-1} + 1} + \ldots + \xi_{\eta_k}, \quad \text{Var}(\zeta_k) = B_k, \quad k = 1, \ldots, s$$

and $L_\gamma = \sum_{j=1}^n E(|\xi_j|^\gamma), \ \gamma \geq 2$, and assume that, for all $k = 1, \ldots, s$,

$$C_1 w^2 \leq \rho_*(B_k) \leq \rho^*(B_k) \leq C_2 w^2,$$  \hfill (6.24)

where $w = (L_\gamma)^{1/\gamma}/\log^* s$, with some positive constants $C_1$ and $C_2$. Suppose the
quantities
\[ \lambda_{k,\gamma} = \sum_{j=n_{k-1}+1}^{\eta_k} E\|\xi_j\|^\gamma, \quad k = 1, \ldots, s, \]
satisfy, for some \( 0 < \epsilon < 1 \) and constant \( C_3 \),
\[ C_3 d^{\gamma/2} s^{\gamma} (\log^* s)^{\gamma + 3} \max_{1 \leq k \leq s} \lambda_{k,\gamma} \leq L_\gamma. \tag{6.25} \]

Then one can construct on a probability space independent random vectors \( X_1, \ldots, X_n \) and a corresponding set of independent Gaussian vectors \( Y_1, \ldots, Y_n \) so that \( (X_j)_{j=1}^n \overset{\mathcal{D}}{=} (\xi_j)_{j=1}^n, \ E(Y_j) = 0, \ \text{Var}(Y_j) = \text{Var}(X_j), 1 \leq j \leq n, \) and for any \( z > 0, \)
\[ P \left( \max_{t \leq n} \left| \sum_{i=1}^{t} X_i - \sum_{i=1}^{t} Y_i \right| \geq z \right) \leq C_* L_\gamma z^{-\gamma}. \]

where \( C_* \) is a constant that depends on \( d, \gamma, C_1, C_2 \) and \( C_3 \).

We need to find a suitable sequence \( \eta_k \) that allows us to get constants \( C_1, C_2 \) in (6.24) and \( C_3 \) in (6.25). There are roughly \( q = n/(2k_0m) \) many \( Y_j \) random variables.

Define
\[ l = \lfloor q^{2/\gamma} / \log^2 q \rfloor. \tag{6.26} \]

To apply Proposition 6.3, we choose the sequence \( \eta_k = kl \) and \( s \asymp q/l \). This choice is justified by proving the following series of propositions.
Proposition 6.4. Recall \( \lambda_* \) and \( A_j \) from (2. B) and (6.11) respectively. There exists a constant \( \delta > 0 \) such that

\[
2(\lambda_* + \delta)k_0m \leq \rho_* (\text{Var}(A_j)) \leq \rho^* (\text{Var}(A_j)) \leq \| A_j \|^2 \leq 2k_0m\Theta^2_{0.2}.
\]

Proposition 6.5. We can get positive constants \( c_1 \) and \( c_2 \) such that for all \( j \),

\[
c_1m \leq \rho_* (\text{Var}(Y^\theta_j)) \leq \rho^* (\text{Var}(Y^\theta_j)) \leq E(|Y^\theta_j|^2) \leq c_2m. \tag{6.27}
\]

Proposition 6.6. For \( l \) in (6.26), there exists constant \( c_3 \) such that,

\[
P \left( \max_{1 \leq t \leq q/l} \left| \text{Var} \left( \sum_{j=(t-1)l}^{t-1} Y^a_j \right) - E \left( \text{Var} \left( \sum_{j=(t-1)l}^{t-1} Y^a_j \right) \right) \right| \geq c_3lm \right) \to 0.
\]

Proposition 6.7. We can get constants \( c_4 \) and \( c_5 \) such that

\[
P (c_4q^{2/\gamma}m \leq (L^a_q)^{2/\gamma} \leq c_5q^{2/\gamma}m) \to 1.
\]

Proposition 6.8. Choose \( \eta_k = kl \) with \( l \) being defined in (6.26). Then we can get \( C_1 \) and \( C_2 \) such that (6.24) is satisfied. Moreover, with \( l \) in (6.26), we can get \( C_3 \) such that (6.25) holds.

Thus, we use Proposition 6.3 to construct \( d \)-variate mean zero normal random vectors \( N^a_j \) and random vectors \( E^a_j \) such that

\[
E^a_j \overset{D}{=} Y^a_j \text{ and } \text{Var}(N^a_j) = \text{Var}(Y^a), \quad 0 \leq j \leq q - 1,
\]
\[
P_a \left( \max_{1 \leq i \leq n} |\Pi_i^a - D_i^a| \geq c_0 z \right) \leq C \frac{L^a_i}{z^\gamma}, \text{ where } \Pi_i^a = \sum_{j=0}^{q_i-1} E_j^a, \quad D_i^a = \sum_{j=0}^{q_i-1} N_j^a \quad (6.28)
\]
and \( C \) is a constant depending on \( \gamma, c_1, \ldots, c_5 \) and \( C_3 \). These constants are free of \( a \). We can create a set \( \mathcal{A} \) with \( P(\mathcal{A}) \to 1 \) so that \( a \in \mathcal{A} \) implies the statements in Proposition 6.7 and Proposition 6.6 hold. Putting \( z = n^{1/r} \) above in (6.28), by Lemma 7.3 and the restriction (4.6), we have, as \( n \to \infty \),
\[
E(L^a_i n^{-\gamma/r}) \leq \frac{q}{n^{\gamma/r}} c_\gamma \max_c E(|\bar{R}_{c,2k_0m}|) = O(n^{1-\gamma/r} m^{\gamma/2-1}) \to 0, \quad (6.29)
\]
using
\[
E(|Y_j(\bar{a}_{2k_0}, \bar{a}_{2k_0+2k_0})|^{\gamma}) \leq c_\gamma \max_c E(|\bar{R}_{c,2k_0m}|) = O(m^{\gamma/2}).
\]
Hence, conditioning on whether \( a \) lies in \( \mathcal{A} \) or not, from (6.29) we obtain,
\[
\max_{1 \leq i \leq n} |\Pi_i^\theta - D_i^\theta| = o_P(n^{1/r}). \quad (6.30)
\]

6.3. Unconditional Gaussian approximation and Regrouping: Here we shall work with the processes \( \Pi_i^\theta, \mu_i^\theta \) and \( D_i^\theta \). Note that, \( V_{j0}(\bar{a}_{2k_0}) \) defined in (6.21) is a function of \( \vartheta \) and might not be positive definite in an uniform fashion. For a constant \( 0 < \delta_* < \lambda_* \), let
\[
V_{j1}(\bar{a}_{2k_0}) = \begin{cases} 
V_{j0}(\bar{a}_{2k_0}) & \text{if } \rho_*(V_{j0}^\vartheta) \geq \delta_* m, \\
(\delta_* m)I_d & \text{otherwise},
\end{cases} \quad (6.31)
\]
which is a positive-definitized version of $V_{j0}(\tilde{a}_{2k0j})$. The following proposition shows that partial sums of $V_{j0}(\tilde{a}_{2k0j})$ and $V_{j1}(\tilde{a}_{2k0j})$ are close to each other.

**Proposition 6.9.** For some $\iota > 0$, we have

$$\max_{i \leq n} E\left(\left| \sum_{j=1}^{\max(1,qi-1)} (V_{j0}(\tilde{a}_{2k0j}) - V_{j1}(\tilde{a}_{2k0j})) \right| \right) = o_P(n^{2/\iota-i}).$$

Henceforth in the sequel we will slightly abuse $\max(1,qi-1) = \max(1,|i/(2k0m)| - 1)$ and simply use $qi-1 = |i/(2k0m)| - 1$ for presentational clarity.

**Proof.** of Proposition 6.9. Recall (6.19) for the definition of $F_{4j+1}(\cdot), F_{4j+2}$ etc. Define

$$F_{21} = \sum_{i=m+1}^{2m} \tilde{X}_i.$$

Define the projection operator $P_i$ by

$$P_i Y = E(Y|F_i) - E(Y|F_{i-1}), \quad Y \in \mathcal{L}_1.$$

For $1 \leq j \leq m$, $\|P_j F_{21}\| \leq \sum_{i=m+1-j}^{m} \delta_{i,2}$. Since $\|E(F_{21}^T|F_m)\|^2 = \sum_{j=1}^{m} \|P_j F_{21}\|^2$, we have

$$|E(F_1(\tilde{a}_0)F_2^T)| = |E(F_1(\tilde{a}_0)F_{21}^T)| = |E(F_1(\tilde{a}_0)E(F_{21}^T|F_m))| \leq \|F_1(\tilde{a}_0)\| \left(\sum_{j=1}^{m} \left(\sum_{i=m+1-j}^{m} \delta_{i,2}\right)^2\right)^{1/2}. \quad (6.32)$$

Under the decay condition on $\Theta_{i,p}$ in (2.3), we have

$$E(|E(F_1(\tilde{a}_0)F_{21}^T)\|^\gamma) = O(m^{\max(\gamma/2,\gamma-\gamma)})$$
We expand the last term of $V_{j0}(a_{2k_0j})$ (see (6.21)). Also note that,

$$|E(F_{4j-2}F_{4j-1}^T) + E(F_{4j-1}F_{4j-2}^T)| \ll m \text{ and } \rho_*(\text{Var}(F_{4j+2})) \geq (k_0 - 1)\lambda_*m.$$ 

Then Proposition 6.9 follows from the fact that our solution of $\gamma$ from (4.5), (4.6), and (4.7) satisfy $\gamma > \max(2, 4\chi)$ for $\chi \leq \chi_0$ and

$$n \max_j P\left(\rho_*(V_{j0}^n) < \delta_*m\right) \leq 2n \max_j P(|E(F_{4j+1}(a_{2k_0j}))F_{4j+2}^T| \geq -\theta m/2) = O(n^{2/\gamma - \iota}) = o(n^2/r),$$

for some $\iota > 0$ since we can choose $\delta_*$ such that $\theta = (k_0 - 1)\lambda_* - \delta_* > 0$. \qed

Recall (6.23) for the definition of $U_j$. By Lemma 7.3 and Jensen’s inequality, we obtain $\max_j \|U_j(\bar{\theta}_{2k_0j+2k_0})\|_{\gamma/2} = O(m^{1/2})$. By (4.6), $\phi_n := q^{1/\gamma}m^{1/2}n^{-1/r} \to 0$. Then

$$P\left(\max_{0 \leq j \leq q-1} |U_j(\bar{\theta}_{2k_0j+2k_0})| \geq \phi_n n^{2/r}\right) \leq \sum_{j=0}^{q-1} P\left(|U_j(\bar{\theta}_{2k_0j+2k_0})| \geq \phi_n n^{2/r}\right) = O(\phi_n^{-\gamma/2} n^{1-\gamma/r} m^{\gamma/2 - 1}) = O(\phi_n^{\gamma/2}) \to 0.$$

Similarly, $|L(\bar{\theta}_0)| = o_P(n^{2/r})$. Thus, by (6.22) and Proposition 6.9, since $\text{Var}(Y_j^a) = \text{Var}(N_j^a)$, one can construct i.i.d. $N(0, I_d)$ normal vectors $Z_l^a, l \in \mathbb{Z}$, such that

$$\max_{i \leq n} |D_i^\theta - \varsigma_i(\vartheta)| = o_P(n^{1/r}), \text{ where } \varsigma_i(a) = \sum_{j=1}^{q-1} V_{j0}^a(a_{2k_0j})^{1/2}Z_j^a.$$

By (6.30), we have

$$\max_{i \leq n} |\Pi_i^\theta - \varsigma_i(\vartheta)| = o_P(n^{1/r}).$$
Let $Z^*_l, l \in \mathbb{Z}$, independent of $(\epsilon_j)_{j \in \mathbb{Z}}$, be i.i.d. $N(0, I_d)$ and define

$$
\Psi_i = \sum_{j=1}^{q_i-1} V_{j1}(\bar{\vartheta}_{2k_0j})^{1/2} Z^*_j.
$$

From the distributional equality,

$$
(\Pi^\vartheta_i + M_i(\vartheta))_{1 \leq i \leq n} \overset{D}{=} (S^\vartheta_i)_{1 \leq i \leq n},
$$

we need to prove Gaussian approximation for the process $\Psi_i + M_i(\vartheta)$. Define

$$
B_j = V_{j1}(\bar{\vartheta}_{2k_0j})^{1/2} Z^*_j + \Lambda_{4j}(\bar{\vartheta}_{2k_0j}) + \Lambda_{4j+1}(\bar{\vartheta}_{2k_0j}),
$$

which are independent random vectors for $j = 1, \ldots, q$ and let

$$
S^\vartheta_i = \sum_{j=1}^{q_i-1} B_j \quad \text{and} \quad W^\vartheta_i = \Psi_i + M_i(\vartheta) - S^\vartheta_i.
$$

Note that,

$$
\max_{i \leq n} |W^\vartheta_i| = \max_{i \leq n} |\Lambda_{4q_i}(\bar{\vartheta}_{2k_0q_i}) + \Lambda_1(\bar{\vartheta}_0)| = O_P(n^{1/r}). \quad (6.34)
$$

Conditions (6.24) and (6.25) can be verified easily with this unconditional process $(S)^\vartheta_i$ to use the Proposition 6.3. Thus, there exists $B^\text{new}_j$ and Gaussian random variable $B^\text{gau}_j$, such that $(B^\text{new}_j)_{j \leq q-1} \overset{D}{=} (B_j)_{j \leq q-1}$ and corresponding $B^\text{gau}_j \sim N(0, Var(B_j))$, such that

$$
\max_{i \leq n} \left| \sum_{j=1}^{[i/2k_0m]-1} B^\text{new}_j - \sum_{j=1}^{[i/2k_0m]-1} B^\text{gau}_j \right| = O_P(n^{1/r}). \quad (6.35)
$$
By (6.16), (6.33), (6.34) and (6.35), we can construct a process $S^c_i$ and $B^c_j$ such that 

$(S_i^c)_{i \leq n} \overset{D}{=} (S_i)_{i \leq n}$ and $(B_j^c)_{j \leq q-1} \overset{D}{=} (B_j^{gauss})_{j \leq q-1}$ and

$$\max_{i \leq n} |S^c_i - \sum_{j=1}^{\lfloor i/(2km) \rfloor - 1} B^c_j| = o_p(n^{1/r}). \quad (6.36)$$

Relabel this final Gaussian process as

$$G_i^c = \sum_{j=1}^{\lfloor i/(2km) \rfloor - 1} (Var(B_j))^{1/2} Y^c_j,$$

where $Y^c_j$ are i.i.d. $N(0, I_d)$. This concludes the proof of Theorem 2.1. \qed

**Proof.** of Proposition 6.4. Without loss of generality, we prove it for $j = 1$. Note that

$$2km\lambda_* \leq \rho_*(Var(S_{2km})) \leq \rho^*(Var(S_{2km})) \leq \| \sum_{i=1}^{2km} X_i \|_2^2 \leq 2km\Theta_0^2. \quad (6.37)$$

Recall $X^\oplus_i$ and $\tilde{X}_i$ from (6.6) and (6.7). The same upper bound works for $S^\oplus_i$ and $\tilde{S}_i$. Note that, $\|S^\oplus_{2km} - S_{2km}\| = o(m)$ and from [14], we have

$$\| A_1 - S^\oplus_{2km} \| = O(\sqrt{2km}\Theta_{m,2}) = o(\sqrt{2km}).$$

This concludes the proof using the Cauchy-Schwartz inequality. \qed

**Proof.** of Proposition 6.5. As $A_j$ is the block sum of the $m$-dependent processes with length $2km$, we have, using (6.37), for all $j$,

$$2km(\lambda_* + \delta) \leq E(|A_j|^2) \leq 2km\Theta_0^2.$$
for some small $\delta > 0$. We conclude the proof by using

$$|E(|Y_j^\delta|^2) - E(|A_{j+1}|^2)| = |\Lambda_{4j+1}(\bar{\gamma}_{2k_0j})|^2 + |\Lambda_{4j+4}(\bar{\gamma}_{2k_0j+2k_0})|^2 \leq 2m\Theta_{0,2}^2$$

and $k_0 > \Theta_{0,2}^2/\lambda_* + 1$. Using similar arguments, (6.27) follows.

**Proof.** of Proposition 6.6. Note that, without loss of generality, we can assume $V_j^a$ to be independent for different $j$ since otherwise we can always break the probability statement in even and odd blocks and prove the statement separately. We use Corollary 1.6 and Corollary 1.7 from Nagaev (1979, [18]) respectively for the case $\gamma < 4$ and $\gamma \geq 4$ on $|V_j^a - E(V_j^a)|$ to deduce that it suffices to show the following

$$\max_{1 \leq t \leq q} \max_{t/(t-1)+1 \leq j \leq q} P(|V_j^a - E(V_j^a)| \geq lm) \to 0. \quad (6.38)$$

We expand and write $V_j^a$ as follows:

$$V_j^a = Var(F_{4j+1}(\bar{a}_{2k_0j}) - \Lambda_{4j+1}(\bar{a}_{2k_0j})) + Var(F_{4j+2} + F_{4j+3})$$

$$+ E((F_{4j+1}(\bar{a}_{2k_0j}) - \Lambda_{4j+1}(\bar{a}_{2k_0j}))F_{4j+2}^T) + E(F_{4j+2}(F_{4j+1}(\bar{a}_{2k_0j}) - \Lambda_{4j+1}(\bar{a}_{2k_0j}))^T)$$

$$+ E(F_{4j+3}(F_{4j+4}(\bar{a}_{2k_0j+2k_0}) - \Lambda_{4j+4}(\bar{a}_{2k_0j+2k_0}))^T)$$

$$+ E((F_{4j+4}(\bar{a}_{2k_0j+2k_0}) - \Lambda_{4j+4}(\bar{a}_{2k_0j+2k_0}))F_{4j+3}^T)$$

$$+ Var(F_{4j+4}(\bar{a}_{2k_0j+2k_0}) - \Lambda_{4j+4}(\bar{a}_{2k_0j+2k_0})).$$

Using derivation similar to (6.32), it suffices to show (6.38) for only the first and last term in (6.39). Moreover, we assume $d = 1$ and $j = 1$ to simplify notations.
The proofs and the theorems used can be easily extended to vector-valued processes. Denote by \( \tilde{S}_{m,j} \) for the sum \( \tilde{S}_m \) with \( \epsilon_j \) replaced by an i.i.d. copy \( \epsilon'_j \). For the first term, by Burkholder’s inequality,

\[
E(\|\text{Var}(F_1(\bar{a}_0)) - E(\text{Var}(F_1(\bar{a}_0)))\|^{\gamma/2}) = E(E(\|\tilde{S}_m^2(a_{1-m}, \ldots, a_0) - \tilde{S}_m^2\|^{\gamma/2})) \\
= \| \sum_{j=-m}^{0} P_j \tilde{S}_m^2 \|^{\gamma/2} \leq c_{\gamma} (\sum_{j=-m}^{0} \| P_j \tilde{S}_m^2 \|^{2\gamma/2})^{\gamma/4}
\]

For \(-m \leq j \leq 0\), \( \| P_j \tilde{S}_m^2 \|^{\gamma/2} \leq \| \tilde{S}_m^2 - \tilde{S}_{m,j}^2 \|^{\gamma/2} \leq \| \tilde{S}_m - \tilde{S}_{m,j} \| \| \tilde{S}_m + \tilde{S}_{m,j} \|^{\gamma} \).

Note that \( \| \tilde{S}_m \|^{\gamma} = O(m^{1/2}) \) and \( \| \tilde{S}_m - \tilde{S}_{m,j} \|^{\gamma} \leq \sum_{r=1}^{m} \tilde{\delta}_{r-j,\gamma} \). By Lemma 7.2, \( \tilde{\delta}_{k,\gamma} \leq 2n^{1-p-1/\gamma} t_n^{1-p/\gamma} \delta_{k,p}^{p/\gamma} \). Then since \( 3 > 2(\chi + 1)p/\gamma \) for \( \chi \leq \chi_0 \), we have

\[
\sum_{j=-m}^{0} \| P_j \tilde{S}_m^2 \|^{\gamma/2} = O(m) \sum_{j=-m}^{0} \sum_{r=1}^{m} (\tilde{\delta}_{r-j,\gamma})^2 \tag{6.40}
\]

\[
= O(m) n^{2(p-2)/\gamma} t_n^{2-2p/\gamma} \sum_{j=0}^{m} (\sum_{r=1}^{m} \delta_{r+j,p}^{p/\gamma})^2
\]

\[
= O(m) n^{2(p-2)/\gamma} t_n^{2-2p/\gamma} m^{3-2(\chi+1)p/\gamma} (\log m)^{-2Ap/\gamma},
\]

by (2.3) and the Hölder inequality. Then, since \( A > 2\gamma/p \) and \( \log m \sim \log q \sim \log n \),

\[
qE(\|\text{Var}(F_1(\bar{a}_0)) - E(\text{Var}(F_1(\bar{a}_0)))\|^{\gamma/2}) \leq qm^{\gamma-(\chi+1)p/2} n^{\gamma/2-p-1/2} t_n^{\gamma/2-p/2} (\log n)^{-Ap/2} = o((lm)^{\gamma/2}), \tag{6.41}
\]

using (6.5), (4.7) and the choice of \( l \) in (6.26). For the last term in (6.39), we view \( E(F_4(\bar{a}_{2k_0})^2) \) as

\[
E(F_4(\bar{a}_{2k_0})^2) = E((\tilde{S}_{2k_0,m} - \tilde{S}_{(2k_0-1)m})^2 | a_{(2k_0-1)m+1, \ldots, a_{2k_0,m}})
\]
and show that it is close to \((\tilde{S}_{2k_0m} - \tilde{S}_{(2k_0-1)m})^2\). Let \(F^m_j = (\varepsilon_j, \ldots, \varepsilon_m)\). Note that,

\[
\|\tilde{S}_m^2 - E(\tilde{S}_m^2|a_m, \ldots, a_1)\|_{\gamma/2}^\gamma \lesssim \left( \sum_{j=-m-1}^{0} \|E(\tilde{S}_m^2|F^m_j) - E(\tilde{S}_m^2|F^m_{j+1})\|_{\gamma/2}^2 \right)^{\gamma/4} \quad (6.42)
\]

\[
\leq cm^{\gamma-(\chi+1)p/2}n^{\gamma/2p-1/2}t_n^{\gamma/2-\gamma/p/2}(\log m)^{-Ap/2} = o(q^{-1}(lnm)^{\gamma/2}),
\]

similar to the derivation in (6.40). By (6.41) and (6.42), it suffices to show that

\[
\frac{n}{m} P(|\tilde{S}_m| \geq \sqrt{lm}) \to 0. \quad (6.43)
\]

Using the Nagaev-type inequality from Wu and Wu (2016, [28]) we obtain

\[
P(|\tilde{S}_m| \geq \sqrt{lm}) \leq C_1 \frac{m^{\max\{1, p(1/2-\chi)\}}}{(lm)^{p/2}} + C_2 \exp(-C_3l), \quad (6.44)
\]

where \(C_1, C_2\) and \(C_3\) depend on \(\chi\) and \(p\). The second term in (6.44) is \(o(m/n)\) since \(e^{-l} \to 0\) very fast. For the first term in (6.44), if \(\chi < 1/2 - 1/p\), then

\[
\frac{n}{m} \frac{m^{p(1/2-\chi)}}{(lm)^{p/2}} = (\log n)^p n^{1-p/\gamma + L(p/\gamma - p\chi - 1)} t_n^{k(p/\gamma - p\chi - 1)} = o(1),
\]

as from (4.7) we have \(1 - p/\gamma + L(p/\gamma - p\chi - 1) = L(p/\gamma - 1)(\chi p + p + 1) < 0\). If \(1/2 - 1/p \leq \chi < \chi_0\) and consequently \(r < p\), then we have, for the first term in (6.44),

\[
\frac{n}{m} \frac{m}{(lm)^{p/2}} = (\log n)^p n^{1/p-1/\gamma + L(1/\gamma - 1/2)} t_n^{k(p/\gamma - p/2)} = o(1), \quad (6.45)
\]

using (6.5), \(r < p\) and the fact that \(r\) satisfy \(1/r - 1/\gamma + L(1/\gamma - 1/2) = 0\). \(\square\)
Proof. of Proposition 6.7. By Lemma 7.3, \( E(L_\gamma^a) \approx qm^{\gamma/2} \). Then it suffices to prove

\[
P(|L_\gamma^a - E(L_\gamma^a)| \geq cqm^{\gamma/2}/\log q) \to 0,
\]

(6.46)

holds for some constant \( c > 0 \). Note that \( E(|Y^a_j|^\gamma) \) are even indices \( j \) (also for odd indices \( j \)). Thus we can prove the statement separately by breaking \( L_\gamma^a \) in sum of even and odd \( E(|Y^a_j|^\gamma) \). Without loss of generality, we assume all \( E(|Y^a_j|^\gamma) \) are independent and proceed. Define \( J_j = (2k_0m)^{-\gamma/2} E(|\tilde{S}_{2k_0m_j} - \tilde{S}_{2k_0m(j-1)}|^\gamma|\tilde{a}_{2k_0(j-1)}, \tilde{a}_{2k_0j}) \) and \( \theta = \ell^{\gamma/2} = q/(\log q)^\gamma \). Recall the truncation operator \( T \) from (6.1). Noting \( E(J_j) = O(1) \) from Lemma 7.3, we have

\[
P(\sum_{j=1}^q T_\theta(J_j) - E(T_\theta(J_j)) \leq \phi) \leq \frac{q}{\phi^2} \max_j E(T_\theta(J_j)^2) = O(\theta q/\phi^2) = o(1),
\]

where \( \phi = q/\log q \), and

\[
\max_j P(J_j \geq \theta) \leq \max_j P(E(|\tilde{S}_{2k_0m_j} - \tilde{S}_{2k_0m(j-1)}|^2|\tilde{a}_{2k_0(j-1)}, \tilde{a}_{2k_0j}) \geq 2k_0lm) = o(q^{-1}),
\]

from (6.41), (6.42) and (6.43). Thus \( P(\sum_{j=1}^q J_j - \sum_{j=1}^q E(J_j) \geq \phi) \to 0 \) which is a restatement of (6.46).

Proof. of Proposition 6.8. We showed in Proposition 6.7 that

\[
P(cqm^{\gamma/2} \leq L_\gamma \leq Cqm^{\gamma/2}) \to 1,
\]
for some constants $c$ and $C$. Let $l$ be as given in (6.26). Let $S = \{0, l, 2l, \cdots \}$.

Proposition 6.5 and Proposition 6.6 show that, for some constants $c$ and $C$,

$$P(\text{clk}_0 m \leq \min_{i \in S} \rho_*(\text{Var} \left( \sum_{j=l}^{i+l-1} Y_j^a \right))) \leq \max_{i \in S} \rho^* \left( \text{Var} \left( \sum_{j=l}^{i+l-1} Y_j^a \right) \right) \leq C \text{clk}_0 m \to 1.$$

We choose $\eta_k = kl$ and $s \gg q/l$. Starting with the conditional block sum process $Y_j^a$ for $0 \leq j \leq q - 1$, this choice of $\eta_k$ satisfies (6.24) for a given $a$ with probability going to 1. The other condition, (6.25) can be easily verified for such a choice of $\eta$-sequence using ideas similar to the proof of Proposition 6.7. We skip the details of that derivation. \qed

7. Some Useful Results

Lemma 7.1. Let $p < \gamma$. Assume (2.A). Then $\sup_i E \min \{|X_i|^{\gamma n^{-\gamma/p}}, 1\} = o(n^{-1})$.

Proof. Choose $k_n = \lceil 2(\log n)/(p+\gamma) \log 2 \rceil$. Then $n = o(2^{\gamma k_n})$ and $2^{pk_n} = o(n)$.

Let $Z = |X_i|n^{-1/p}$. The lemma follows from

$$E(\min\{Z^\gamma, 1\}) \leq P(Z \geq 1) + \sum_{k=0}^{k_n} 2^{-k\gamma} P(2^{-1-k} \leq Z < 2^{-k}) + 2^{-\gamma(k_n+1)}$$

$$\leq E(Z^p 1_{Z \geq 1}) + \sum_{k=0}^{k_n} 2^{p(k+1)-k\gamma} E(Z^p 1_{Z \geq 2^{-1-k}}) + 2^{-\gamma(k_n+1)} = o(n^{-1}),$$

in view of the uniform integrability condition (2.A) and $n^{1/2}/2^{k_n} \to \infty$. \qed

Lemma 7.2. The functional dependence measures defined on the truncated process $(X_i^{\oplus})$ and the $m$-dependent process $(\bar{X}_i)$, satisfy $\delta_{j,\gamma}^{\oplus} \leq \delta_{j,\gamma}^{\bar{X}} \leq 2n^{1/p-1/\gamma} t_n^{-p/\gamma} \delta_{j,p}^{\bar{X}}$. 

Proof. Since the truncation operator \( T \) is Lipschitz continuous,

\[
(\delta_{j,\gamma}^\oplus)^\gamma = \sup_i E( |T_{i,n^{1/p}}(X_i) - T_{i,n^{1/p}}(X_{i,(i-j)})|^\gamma )
\]

\[
= n^{\gamma/p} t_n^\gamma \sup_i E \left( \min \left( 2, \frac{|X_i - X_{i,(i-j)}|}{t_n^{1/p}} \right)^\gamma \right) \leq 2^{\gamma} n^{\gamma/p - 1} t_n^{\gamma - p} \delta_{j,\gamma}^p.
\]

The first inequality \( \tilde{\delta}_{j,\gamma} \leq \delta_{j,\gamma}^\oplus \) follows from (6.14).

Lemma 7.3. Rosenthal Type Moment Bound Recall (6.4) and (6.5) for \( t_n \). Assume (6.9), (6.12), (6.13) along with (2.6) on \( A \) related to the restriction on \( \Theta_{i,p} \) as mentioned in (2.3). Moreover, assume \( m = \lfloor n^{Lk} \rfloor \) with \( k \) satisfying \( k < (\gamma/2 - 1)^{-1}(\gamma - p) \). Then, we have

\[
\max_t E(\max_{1 \leq l \leq m} |\tilde{R}_{t,l}|^\gamma) = O(m^{\gamma/2}). \tag{7.1}
\]

Proof. Since the functional dependence measure is defined in an uniform manner, we can ignore the \( \max_t \) in (7.1) and use the Rosenthal-type inequality for stationary processes in Liu, Xiao and Wu (2013, [15]). By [15], there is a constant \( c \), depending only on \( \gamma \), such that

\[
\| \max_{1 \leq l \leq m} |\tilde{R}_{t,l}| \|_\gamma \leq cm^{1/2} \left[ \sum_{j=1}^{m} \tilde{\delta}_{j,\gamma} + \sum_{j=1+m}^{\infty} \tilde{\delta}_{j,\gamma} + \sup_i \|T_{i,n^{1/p}}(X_i)\| \right]
\]

\[
+ cm^{1/\gamma} \left[ \sum_{j=1}^{m} j^{1/2 - 1/\gamma} \tilde{\delta}_{j,\gamma} + \sup_i \|T_{i,n^{1/p}}(X_i)\|_{\gamma} \right]
\]

\[
\leq c(I + II + III + IV),
\]
where

\[ I = m^{1/2} \sum_{j=1}^{m} \tilde{\delta}_{j,2} + m^{1/2} \|X_1\|_2, \]

\[ II = m^{1/2} \sum_{j=m+1}^{\infty} \tilde{\delta}_{j,\gamma}, \quad III = m^{1/\gamma} \sum_{j=1}^{\infty} j^{1/2 - 1/\gamma} \tilde{\delta}_{j,\gamma}, \]

\[ IV = m^{1/\gamma} \sup_t \|T_{t_n^{1/p}}(X_t)\|_{\gamma}. \]

For the first term \( I \), since \( \sum_{j=1}^{\infty} \delta_{j,2} \leq \|X_i\|_2 \leq 2\Theta_{0,2} \) and \( \tilde{\delta}_{j,2} \leq \delta_{j,2} \), we have \( I = O(m^{1/2}) \). Starting with \( II \), we apply Lemma 7.2 to obtain

\[ II = m^{1/2} \sum_{j=m+1}^{\infty} \tilde{\delta}_{j,\gamma} \lesssim m^{1/2} n^{(1-p)/(1-p/\gamma)} \sum_{j=m+1}^{\infty} \tilde{\delta}_{j,p/\gamma}. \]

The rest follows from the derivation in (4.4) and (4.7). For the third term, we have

\[ III \lesssim m^{1/\gamma} n^{(1-p/\gamma) - \gamma t_n^{1-p/\gamma} \sum_{j=1}^{m} j^{1/2 - 1/\gamma} \tilde{\delta}_{j,p/\gamma}. \]

\[ \leq m^{1/\gamma} n^{(1-p/\gamma) - \gamma t_n^{1-p/\gamma}} \sum_{l=1}^{[\log_2 m]+1} 2^{l-1} \sum_{j=2^{l-1}}^{\gamma t_n^{1-p/\gamma}} j^{1/2 - 1/\gamma} \tilde{\delta}_{j,p/\gamma}. \]

\[ \leq m^{1/\gamma} n^{(1-p/\gamma) - \gamma t_n^{1-p/\gamma}} \sum_{l=1}^{[\log_2 m]+1} 2^{(3/2 - 1/\gamma - p/\gamma) O}\left(2^{-(l/p)}\gamma 1^{-Ap/\gamma}\right). \]

Recall the definition of \( \chi_0 \) from (2.5). If \( \chi \leq \chi_0 \), then our solution for \( \gamma \) satisfies

\[ 3/2 - 1/\gamma - (\chi + 1)p/\gamma \geq 0, \]

with equality holding only for \( \chi = \chi_0 \). Hence, if \( \chi < \chi_0 \), we have

\[ m^{-1/2} III = m^{1-(\chi+1)p/\gamma} n^{1/p - 1/\gamma t_n^{1-p/\gamma}} (\log n)^{-Ap/\gamma} O(1) = o(1), \]
from (4.7), (6.15) and (6.5). If \( \chi = \chi_0 \), since \( A > \gamma/p \) from (2.6) [The lower bound for \( A \) there is just \( 2\gamma/p \) as mentioned in (4.5)], we have

\[
m^{-1/2}III = m^{1/\gamma-1/2}n^{1/p-1/\gamma}t_n^{1-p/\gamma}O(1) = o(1),
\]

since (4.6) is true. Also for the case of \( \chi > \chi_0 \) in the proof of Theorem 2.2, the way we define our three conditions in (5.1) the new solution also satisfy \( \gamma' = 2(1+p+p\chi)/3 \) and thus (7.3) holds. For the fourth term \( IV \), we use (6.4) to derive

\[
m^{-\gamma/2}IV^\gamma = m^{-\gamma/2} \sup_i \| T_{t_n^{1/p}}(X_i) \| \cdot \gamma
\]

\[
\leq m^{-\gamma/2} t_n^{\gamma/p} \sup_i E \left( \min\left\{ \frac{|X_i|^\gamma}{t_n^{\gamma/p}}, 1 \right\} \right)
\]

\[
= m^{-\gamma/2} t_n^{\gamma/p} o(1) = o(1),
\]

in the light of (4.6). \( \square \)