A Proofs

A.1 Auxiliary results

The following section lists all auxiliary Lemmas required to prove main results. We had to extract common parts from the proof of main Theorems in order to make them readable. Therefore we recommend reading Lemma 2 and the note about notation underneath it first, then main proofs from the sections A.2 and A.3 and finally this list of Lemmas.

Proposition 4. *[18, p.259, Equation 2.1] If process $\{Z_t, \mathcal{F}_t\}_{t\in\mathbb{N}}$ is τ-dependent and $\mathcal{F}$ is rich enough (see [8, Lemma 5.3]), then there exists, for all $t < t_1 < ... < t_l$, $l \in \mathbb{N}$, a random vector $(Z^*_{t_1}, ..., Z^*_{t_l})$ that is independent of $\mathcal{F}_t$, has the same distribution as $(Z_{t_1}, ..., Z_{t_l})$ and*

$$\mathcal{E}\|(Z^*_{t_1}, ..., Z^*_{t_l}) - (Z_{t_1}, ..., Z_{t_l})\|_1 \leq l\tau(t_1 - t).$$

Lemma 1. *Let $\{Z_i, \mathcal{F}_i\}$ be a τ-mixing sequence,$\{\delta_i\}$ a sequence of $i.i.d$ random variables independent of filtration $\mathcal{F}$, a non-deceasing sequence $(i_1 \leq ... \leq i_m)$, a positive integer k such that $1 < k < m$ and some random vector $(Z_{i_1}, ..., Z_{i_m})$. Further let $A = (Z_{i_1}, ..., Z_{i_{k-1}})$, $B = Z_{i_k}$ and $C = (Z_{i_{k+1}}, ..., Z_{i_m})$, $\mathcal{F}_\mathcal{A} = \mathcal{F}_{k-1}$, $\mathcal{F}_\mathcal{B} = \mathcal{F}_k$. There exist independent random variables B^* and C^*, independent of $\mathcal{F}_\mathcal{A}$, such that*

$$\mathcal{E}|B - B^*| = \tau(i_k - i_{k+1}) \text{ and } \frac{1}{m-k}\mathcal{E} \parallel C - C^* \parallel_1 \leq \tau(i_{k+1} - i_k) \tag{12}$$

Proof. We first use use [18, Equation 2.1]] (also [8, Lemma 5.3]) to construct C^* such that $\frac{1}{m-k}\mathcal{E} \parallel C - C^* \parallel_1 \leq (m-k)\tau(i_{k+1} - i_k)$. By construction C^* is independent of $\mathcal{F}_\mathcal{B}$. Since $\mathcal{F}_\mathcal{A} \subset \mathcal{F}_\mathcal{B}$ and $\sigma(B) \subset \mathcal{F}_\mathcal{B}$ and $C^* \perp\!\!\!\perp (\sigma(\delta_k))$, then $C^* \perp\!\!\!\perp (\mathcal{F}_\mathcal{A} \vee \sigma(B) \vee \sigma(\delta_k))$. Next by [8, Lemma 5.2] we construct B^* such that $\mathcal{E}|B - B^*| = \tau(i_k - i_{k+1})$, and B^* independent of $\mathcal{F}_\mathcal{A}$ but $\mathcal{F}_\mathcal{A} \vee \sigma(B) \vee \sigma(\delta_k)$ measurable. Since $\sigma(C^*) \perp\!\!\!\perp (\mathcal{F}_\mathcal{A} \vee \sigma(B) \vee \sigma(\delta))$ then C^* and B^* are independent. Finally both C^* and B^* are independent of $\mathcal{F}_\mathcal{A}$ □

Lemma 2. *[24, Section 5.1.5] Any core h can be written as a sum of canonical cores $h_1, ..., h_m$ and a constant h_0*

$$h(z_1, ..., z_m) = h_m(z_1, ..., z_m) + \sum_{1 \leq i_1 < ... < i_{m-1} \leq m} h_{m-1}(z_{i_1}, ..., z_{i_{m-1}})$$
$$+ ... + \sum_{1 \leq i_1 < i_2 \leq m} h_2(z_{i_1}, z_{i_2}) + \sum_{1 \leq i \leq m} h_1(z_i) + h_0$$

Proof. To show this we define auxiliary functions

$$g_c(z_1, ...z_c) = \mathcal{E}h(z_1, ..., z_c, Z^*_{c+1}, ..., Z^*_m)$$

for each $c = 0, ..., m-1$ and put $g_m = h$.

Canonical functions that allow core decomposition are

$$h_0 = g_0, \tag{13}$$
$$h_1(z_1) = g_1(z_1) - h_0, \tag{14}$$
$$h_2(z_1, z_2) = g_2(z_1, z_2) - h_1(z_1) - h_1(z_2) - h_0, \tag{15}$$
$$h_3(z_1, z_2, z_3) = g_3(z_1, z_2, z_3) - \sum_{1 \leq i < j \leq 3} h_2(z_i, z_j) - \sum_{1 \leq i \leq 3} h_1(z_i) - h_0, \tag{16}$$
$$\cdots, \tag{17}$$
$$h_m(z_1, ..., z_m) = g_m(z_1, ..., z_m) - \sum_{1 \leq i_1 < ... < i_{m-1} \leq m} h_{m-1}(z_{i_1}, ..., z_{i_{m-1}}) \tag{18}$$
$$- ... - \sum_{1 \leq i_1 < i_2 \leq m} h_2(z_{i_1}, z_{i_2}) - \sum_{1 \leq i \leq m} h_1(z_i) - h_0. \tag{19}$$
$$\tag{20}$$

Lemma 3 shows that functions h_c are symmetric (and therefore cores) and Lemma 4 shows that they are canonical. □

We call $h_1, ..., h_m$ components of a core h. We do not call h_0 a component, its simply a constant.

Lemma 3. *[24, Section 5.1.5] Components of a core h are symmetric functions.*

Lemma 4. *[24, Section 5.1.5] A component of a core h is a canonical core.*

Lemma 5. *If h is bounded and Lipschitz continuous core then its components are also bounded and Lipschitz continuous.*

Proof. Note that

$$g_c(z_1, ...z_c) = \mathcal{E}h(z_1, ..., z_c, Z^*_{c+1}, ..., Z^*_m) \leq \mathcal{E} \parallel h \parallel_\infty . \tag{21}$$

To prove boundedness we use induction - we assume that components with low index are bounded and use the fact that sum of bounded functions is bounded to obtain the required results. We prove Lipschitz continuity similarly, first by showing that $g_c(z_1, ...z_c)$ are Lipschitz continuous with the same coefficient as the core h and then by using the fact that sum of Lipschitz continuous functions is Lipschitz continuous. □

Lemma 6. *If $1 \leq j_k, r_j \leq m$ are disjoint sequences with respectively q and $m - q$ elements, such that elements in each sequence are unique then*

$$\sum_{i \in N^m} f(Z_{i_{j_1}}, ..., Z_{i_{j_q}}) = n^{m-q} \sum_{i \in N^q} f(Z_{i_1}, ..., Z_{i_q}) \tag{22}$$

Proof.

$$\sum_{i \in N^m} f(Z_{i_{j_1}}, ..., Z_{i_{j_q}}) = \sum_{1 \leq i_{j_1}, ..., i_{j_q} \leq n} \sum_{1 \leq i_{r_1}, ..., i_{r_{m-q}} \leq n} f(Z_{i_{j_1}}, ..., Z_{i_{j_q}}) = \tag{23}$$

$$\sum_{1 \leq i_{j_1}, ..., i_{j_q} \leq n} \left(f(Z_{i_{j_1}}, ..., Z_{i_{j_q}}) \sum_{1 \leq i_{r_1}, ..., i_{r_{m-q}} \leq n} 1 \right) = \tag{24}$$

$$n^{m-q} \sum_{1 \leq i_{j_1}, ..., i_{j_q} \leq n} f(Z_{i_{j_1}}, ..., Z_{i_{j_q}}) = n^{m-q} \sum_{i \in N^q} f(Z_{i_1}, ..., Z_{i_q}). \tag{25}$$

□

Lemma 7. *[Section 5.1.5] V-statistic of a core function h can be written as a sum of V-statistics with canonical cores*

$$V(h) = V(h_m) + \binom{m}{1} V(h_{m-1}) + ... + \binom{m}{m-2} V(h_2) + \binom{m}{m-1} V(h_1) + h_0. \tag{26}$$

Lemma 8. *Assume that the stationary process Z_t is τ-dependent with a coefficient $\tau(i) = i^{-6-\epsilon}$ for some $\epsilon > 0$. If h is a canonical and Lipschitz continuous core of three arguments then*

$$\lim_{n \to \infty} \frac{1}{n^2} \sum_{i \in N^m} |\mathcal{E}h(Z_{i_1}, Z_{i_2}, Z_{i_3})| = 0. \tag{27}$$

converges to zero in probability.

Proof. We change summing order, it is useful to think of index b as 'beginning', e as 'end' and m as 'middle'.

$$\sum_{i \in N^m} |\mathcal{E}h(Z_{i_1}, Z_{i_2}, Z_{i_3})| = 3! \sum_{b=1}^{n} \sum_{e=b}^{n} \sum_{b \leq m \leq e}^{n} |\mathcal{E}h(Z_b, Z_m, Z_e)|. \tag{28}$$

For each b, m, e, $|\mathcal{E}h(Z_b, Z_m, Z_e)| \leq 2\tau(max(e - m, m - b))$. To see that suppose that $m - b > e - m$. Then by Proposition 4 there exists random vector (Z^*_m, Z^*_e) independent of Z_b such that $\frac{1}{2}\mathcal{E} \parallel (Z^*_m, Z^*_e) - (Z_m, Z_e) \parallel \leq \tau(m - b)$. Furthermore

$$|\mathcal{E}\left(h(Z_b, Z_m, Z_e) - h(Z_b, Z^*_m, Z^*_e) + h(Z_b, Z^*_m, Z^*_e)\right)| \leq \tag{29}$$

$$|\mathcal{E}\left(h(Z_b, Z_m, Z_e) - h(Z_b, Z^*_m, Z^*_e)\right)| + |\mathcal{E}h(Z_b, Z^*_m, Z^*_e)| \leq Lip(h)2\tau(m - b) + 0, \tag{30}$$

since $\mathcal{E}h(Z_b, Z_m^*, Z_e^*) = 0$. Similar reasoning for case $m - b < e - m$ proofs that $|\mathcal{E}h(Z_b, Z_m, Z_e)| \leq 2Lip(h)\tau(max(e-m, m-b))$. Since $max(e-m, m-b) > (e-b)/2$

$$3!\sum_{b=1}^{n}\sum_{e=b}^{n}\sum_{b\leq m\leq e}^{n}|\mathcal{E}h(Z_b, Z_m, Z_e)| \leq 12Lip(h)\sum_{b=1}^{n}\sum_{e=b}^{n}\sum_{b\leq m\leq e}^{n}\tau((e-b)/2) \leq \tag{31}$$

$$12Lip(h)\sum_{b=1}^{n}\sum_{e=b}^{n}(e-b)\frac{8}{(e-b)^3} \leq 96Lip(h)\sum_{b=1}^{n}\sum_{e=b}^{n}\frac{1}{(e-b)^2} = O(n). \tag{32}$$

□

Lemma 9. *Assume that the stationary process Z_t is τ-dependent with a coefficient $\tau(i) = i^{-6-\epsilon}$ for some $\epsilon > 0$. If h is Lipschitz continuous function of m arguments, $m > 3$, such that for any $1 \leq k \leq m$*

$$\mathcal{E}h(z_1, ..., Z_k, ..., z_m) = 0 \tag{33}$$

then

$$\lim_{n\to\infty}\frac{1}{n^{m-2}}\sum_{i\in N^m}|\mathcal{E}h(Z_{i_1}, ..., Z_{i_m})| = 0. \tag{34}$$

converges to zero in probability.

Proof. The proof follows proof by [1, Lemma 3].

$$\sum_{i\in N^m}|\mathcal{E}h(Z_{i_1}, ..., Z_{i_m})| \leq \sum_{\pi\in S_m}\sum_{1\leq i_1<...\leq i_m\leq n}|\mathcal{E}h(Z_{i_{\pi(1)}}, ..., Z_{i_{\pi(m)}})| \tag{35}$$

where S_m is group of permutations of m elements. Let, $g = \lfloor m/2 \rfloor$, $j_1 = i_2 - i_1$, let $j_l = min(i_{2l-1} - i_{2l-2}, i_{2l} - i_{2l-1})$ for $2 \leq l \leq g$ and let $j_g = i_m - i_{m-1}$ if m is even. If j_1 is equal to $max(j_1, ..., j_g)$ then we use Proposition 4 and, by the reasoning similar to one in the Lemma 8), we obtain bound

$$|\mathcal{E}h(Z_{i_{\pi(1)}}, ..., Z_{i_{\pi(m)}})| \leq \tau(j_1). \tag{36}$$

Same reasoning holds if j_g is equal to $max(j_1, ..., j_g)$ and m is even. In other case there exists $1 < k \leq g$ for which maximum is obtained. Let

$$A = (Z_{i_1}, ..., Z_{i_{2k-1}}) \tag{37}$$
$$B = Z_{i_{2k}} \tag{38}$$
$$C = (Z_{i_{2k+1}}, ..., Z_{i_m})) \tag{39}$$
$$h(A, B, C) = h(Z_{i_{\pi(1)}}, ..., Z_{i_{\pi(2k-1)}}, Z_{i_{\pi(2k)}}, Z_{i_{\pi(2k+1)}}, ..., Z_{i_\pi(m)}). \tag{40}$$

And B^*,C are as in Lemma 1. We use Lemma 1 to see that

$$|\mathcal{E}h(A, B, C)| \leq |\mathcal{E}(h(A, B, C) - h(A, B^*, C^*)| + |\mathcal{E}h(A, B^*, C^*)| = \tag{41}$$
$$= |\mathcal{E}(h(A, B, C) - h(A, B^*, C^*)| + 0 \leq Lip(h)\tau(j_k). \tag{42}$$

For second equality we have used assumption 33 and that B^* is independent of A and C^*. Therefore we showed that if $w = max(j_1, ..., j_g)$

$$|\mathcal{E}h(Z_{i_{\pi(1)}}, ..., Z_{i_{\pi(m)}})| \leq \tau(w). \tag{43}$$

Now for each $w = 1$ to n we count all possible combinations of indexes i_1 to i_m. Suppose $w = max(j_1, ..., j_g)$ and that it is obtained at first position i.e. $j_1 = w$. Then i_1 can take at most n positions and position of i_2 is fixed. If $i_3 - i_2 \leq i_4 - i_3$ then $i_3 \leq w + i_2$ and therefore i_3 can take at most w values and i_4 can take at most n values. On the other hand if $i_3 - i_2 \geq i_4 - i_3$ then i_4 can take at most w values and i_3 can take at most n values. Proceeding in this way we obtain that the possible values for the variables $i_1 \leq ... \leq i_m$. In case m is even number of combinations of $i_1, ..., i_m$ is smaller than $n^g w^{g-1}$ and in case m is odd number of combinations is smaller than $n^g w^g$.

If $j_k = max(j_1, ..., j_g)$ for $1 < k \leq g$ similar reasoning holds - we bound number of combinations for the first triple (i_1, i_2, i_3), second triple (i_3, i_4, i_5) etc.

There are $O(n)$ combinations for $w = 0$ and since $\parallel h \parallel_\infty \leq \infty$ sum over them is of order n. Therefore for values of w varying in a range 1 to $n - 1$ we have

$$\sum_{\pi \in S_m} \sum_{1 \leq i_1 < \ldots \leq i_m \leq n} |\mathcal{E}h(Z_{i_{\pi(1)}}, ..., Z_{i_{\pi(m)}})| \leq \tag{44}$$

$$n^g \left(\sum_{w=1}^{n-1} w^g \tau(g) \right) + O(n) \leq n^g \left(\sum_{w=1}^{n-1} w^{g-4} \right) + O(n) = O(n^{2g-3}) = O(n^{m-3}). \tag{45}$$

□

Corollary 1. *Assume that the stationary process Z_t is τ-dependent with a coefficient $\tau(i) = i^{-6-\epsilon}$ for some $\epsilon > 0$. If h is a canonical and Lipschitz continuous core of three or more arguments, then*

$$\lim_{n \to \infty} \frac{1}{n^{m-1}} \sum_{i \in N^m} |\mathcal{E}h(Z_{i_1}, ..., Z_{i_m})| = 0. \tag{46}$$

Lemma 10. *Assume that the stationary process Z_t is τ-dependent with a coefficient $\tau(i) = i^{-6-\epsilon}$ for some $\epsilon > 0$. If h is a canonical and Lipschitz continuous core of three or more arguments, then*

$$\lim_{n \to \infty} \frac{1}{n^{2m-2}} \sum_{i \in N^{2m}} \mathcal{E}|h(Z_{i_1}, ..., Z_{i_m})h(Z_{i_{m+1}}, ..., Z_{i_{2m}})| = 0.$$

Proof. Let $g(z_{i_1}, ..., z_{i_m}, z_{i_{m+1}}, ..., z_{i_{2m}}) = h(z_{i_1}, ..., z_{i_m})h(z_{i_{m+1}}, ..., z_{i_{2m}})$. Since h is canonical g meets assumptions of the Lemma 9 from which the Proposition follows. □

Lemma 11. *Assume that the stationary process Z_t is τ-dependent with a coefficient $\tau(i) = i^{-6-\epsilon}$ for some $\epsilon > 0$. Let h be a canonical and Lipschitz continuous core of c arguments, $3 \leq c \leq m$, and $1 \leq j_1 < \ldots < j_c \leq m$ be a sequence of c integers. If $Q_{i_1,...,i_m}$ is a random variable independent of $(Z_{i_1}, ..., Z_{i_m})$ such that $\sup_{i \in N^m} \mathcal{E}|Q_i| \leq \infty$ and $\sup_{i \in N^m} \sup_{o \in N^m} \mathcal{E}|Q_i Q_o| \leq \infty$ then*

$$\lim_{n \to \infty} \frac{1}{n^{m-1}} \mathcal{E} \sum_{i \in N^m} Q_i h(Z_{i_{j_1}}, ..., Z_{i_{j_c}}) \stackrel{P}{=} 0. \tag{47}$$

$$\lim_{n \to \infty} \frac{1}{n^{2m-2}} \mathcal{E} \sum_{i \in N^m} \sum_{o \in N^m} Q_i Q_o h(Z_{i_{j_1}}, ..., Z_{i_{j_c}}) h(Z_{i_{o_1}}, ..., Z_{i_{o_c}}) \stackrel{P}{=} 0. \tag{48}$$

$$\tag{49}$$

For the first limit notice that

$$\frac{1}{n^{m-1}} \mathcal{E} \sum_{i \in N^m} Q_i h(Z_{i_{j_1}}, ..., Z_{i_{j_c}}) \leq \frac{1}{n^{m-1}} \sum_{i \in N^m} |\mathcal{E}Q_i||\mathcal{E}h(Z_{i_{j_1}}, ..., Z_{i_{j_c}})| \stackrel{6}{\leq} \tag{50}$$

$$\sup_{i \in N^m} |\mathcal{E}Q_i| \frac{1}{n^{c-1}} \sum_{i \in N^c} |\mathcal{E}h(Z_{i_1}, ..., Z_{i_c})| \xrightarrow{1} 0 \textit{ in probablity}. \tag{51}$$

Similar reasoning, which uses Lemma 10 instead of 1, shows convergence of the second limit.

Lemma 12. *Assume that the stationary process Z_t is τ-dependent with a coefficient $\tau(i) = i^{-6-\epsilon}$ for some $\epsilon > 0$. Let h be a canonical and Lipschitz continuous core of c arguments, $3 \leq c \leq m$. If $Q_{i_1,...,i_m}$ is a random variable independent of $(Z_{i_1}, ..., Z_{i_m})$ such that $\sup_{i \in N^m} \mathcal{E}|Q_i| \leq \infty$ and $\sup_{i \in N^m} \sup_{o \in N^m} \mathcal{E}|Q_i Q_o| \leq \infty$ then*

$$\lim_{n \to \infty} \frac{1}{n^{m-1}} \sum_{i \in N^m} \sum_{1 \leq j_1 < \ldots < j_c < m} \mathcal{E}Q_i h_c(Z_{i_{j_1}}, ..., Z_{i_{j_c}}) \stackrel{P}{=} 0. \tag{52}$$

Proof. For each sequence such that $1 \leq j_1 < \ldots < j_c < m$ we apply Lemma 11 and conclude that the random sum

$$\frac{1}{n^{m-1}} \sum_{i \in N^m} \mathcal{E}Q_{i_{j_1},...,i_{j_c}} h_c(Z_{i_{j_1}}, ..., Z_{i_{j_c}}) \tag{53}$$

converges to zero in a probability - from this the proposition follows. □

Lemma 13. *Assume that the stationary process Z_t is τ-dependent with a coefficient $\tau(i) = i^{-6-\epsilon}$ for some $\epsilon > 0$. If h if a canonical and Lipschitz continuous core of three or more arguments*

$$\lim_{n\to\infty} nV(h_c) = 0. \tag{54}$$

Proof. For each c put $Q = 1$ and use Lemma 12. □

Lemma 14. *If W_i is a bootstrap process defined in the section 2 then*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} W_i \overset{P}{=} 0. \tag{55}$$

Proof. By the definition of W_i, $\mathcal{E}(\sum_{i=1}^{n} W_i)^2 = O(nl_n)$, $\lim_{n\to\infty} \frac{l_n}{n} = 0$ and $\mathcal{E}\sum_{i=1}^{n} W_i = 0$. Therefore $\frac{1}{n}\sum_{i=1}^{n} W_i$ converges to zero in probability. □

Lemma 15. *Assume that the stationary process Z_t is τ-dependent with a coefficient $\tau(i) = i^{-6-\epsilon}$ for some $\epsilon > 0$ and W_i is a bootstrap process defined in the section 2. Let f be a one-degenerate, Lipschitz continuous, bounded core of at least m arguments, $m \geq 2$. Further assume that f_2 is a kernel. Then for a positive integer p*

$$\lim_{n\to\infty} nV(f)\left(\frac{1}{n}\sum_{i=1}^{n} W_i\right)^p \overset{P}{=} 0 \tag{56}$$

Proof. By the Lemma 14 $\frac{1}{n}\sum_{i=1}^{n} W_i$ converges to zero in probability. By Theorem 1 $\frac{1}{n^{m-1}}\sum_{i\in N^m} f(Z_{i_m}, ..., Z_{i_m})$ converges to some random variable. □

Lemma 16. *Assume that the stationary process Z_t is τ-dependent with a coefficient $\tau(i) = i^{-6-\epsilon}$ for some $\epsilon > 0$ and W_i is a bootstrap process defined in the section 2. If h is a Lipschitz continuous, degenerate and bounded core of two arguments then*

$$\frac{1}{n}\sum_{i\in N^2} W_{i_1} h(Z_{i_1}, Z_{i_2}) \tag{57}$$

converges in distribution to some random variable.

Proof.

$$\frac{1}{n}\sum_{i\in N^2} W_{i_1} h(Z_{i_1}, Z_{i_3}) = \frac{1}{4}(V_- + V_+) \text{ where,} \tag{58}$$

$$V_- = n^{-1}\sum_{i\in N^2}(W_{i_1} - 1)h(Z_{i_1}, Z_{i_2})(W_{i_2} - 1), \tag{59}$$

$$V_+ = n^{-1}\sum_{i\in N^2}(W_{i_1} + 1)h(Z_{i_1}, Z_{i_2})(W_{i_2} + 1), \tag{60}$$

$$\tag{61}$$

are normalized V statistics that converge. To see that we use Lemma 17 with $g_+(x) = x + 1$ and $g_-(x) = x - 1$ respectively. The only non-trivial assumption is that $\mathcal{E}|g_+(W_i)|^k < \infty$ and $\mathcal{E}|g_-(W_i)|^k < \infty$ - this follows from $\mathcal{E}|W_i|^k$. □

Lemma 17. *Assume that the stationary process Z_t is τ-dependent with a coefficient $\tau(i) = i^{-6-\epsilon}$ for some $\epsilon > 0$ and W_i is a bootstrap process defined in the section 2. Let $x = (w, z)$ and suppose $f(x_1, x_2) = g(w_1)g(w_2)h(z_1, z_2)$ where g is Lipschitz continuous, $\mathcal{E}|g(W_1)|^3 \leq \infty$ and h is symmetric, Lipschitz continuous, degenerate and bounded. Then*

$$nV(f) = \frac{1}{n}\sum_{i,j} f(X_i, X_j) = \frac{1}{n}\sum_{i,j} g(W_i)g(W_j)h(Z_i, Z_j) \tag{62}$$

converges to some random variable in law.

Proof. We use [17, Theorem 2.1] to show that $nV(f)$ converges that requires checking assumptions *A1 - A3*

Assumption A1. Point (i) requires that the process (W_n, Z_n) is a strictly stationary sequence of $\mathbb{R}^d$-values integrable random variables - this follows from the assumptions. For the point (ii) we put $\delta = \frac{1}{3}$ and check

$$\sum_{r=1}^{\infty} r\tau(r)^{\delta} \leq \sum_{r=1}^{\infty} r r^{-6\frac{1}{3}} = \sum_{r=1}^{\infty} r^{-2} < \infty. \tag{63}$$

Assumption A2. Point (i) requires that the function f is symmetric, measurable and degenerate. Symmetry and measurability are obvious and for the degeneracy we calculate

$$\mathcal{E}g(W_1)g(w)h(Z_1, z) = \mathcal{E}g(W_1)g(w)\mathcal{E}h(Z_1, z) = 0. \tag{64}$$

Point (ii) requires that for $\nu > (2-\delta)/(1-\delta) = 2.5$ (since we have chosen $\delta = \frac{1}{3}$).

$$\sup_{k\in\mathbf{N}} \mathcal{E}|f(X_1, X_k)|^{\nu} < \infty \text{ and } \sup_{k\in\mathbf{N}} \mathcal{E}|f(X_1, X_k^*)|^{\nu} < \infty \tag{65}$$

Both requirements are met since h is bounded and the process $\mathcal{E}|g(W_i)|^3 \leq \infty$.

Assumption A3. Function f is Lipschitz continuous - this is met since both g and h are Lipschitz continuous. □

Lemma 18. *If W_i is a bootstrap process defined in the section 2 then*

$$\sum_{i=1}^{n} \tilde{W}_i = \sum_{i=1}^{n} \left(W_i - \frac{1}{n} \sum_{i=j} W_j \right) = 0. \tag{66}$$

Lemma 19. *Assume that the stationary process Z_t is τ-dependent with a coefficient $\tau(i) = i^{-6-\epsilon}$ for some $\epsilon > 0$ and W_i is a bootstrap process defined in the section 2. If f is canonical, Lipschitz continuous, bounded core then a random variable*

$$\frac{1}{n} \sum_{1\leq i,j\leq n} \tilde{W}_i \tilde{W}_j f(Z_i, Z_j) \tag{67}$$

converges in law.

Proof.

$$\frac{1}{n} \sum_{1\leq i,j\leq n} \tilde{W}_i \tilde{W}_j f(Z_i, Z_j) = \frac{1}{n} \sum_{1\leq i,j\leq n} \left(W_i - \sum_{a=1}^{n} W_a \right) \left(W_i - \sum_{b=1}^{n} W_b \right) f(Z_i, Z_j) = \tag{68}$$

$$\frac{1}{n} \sum_{1\leq i,j\leq n} W_i W_j f(Z_i, Z_j) - \left(\frac{2}{n} \sum_{1\leq i,j\leq n} f(Z_i, Z_j) \right) \left(\frac{1}{n} \sum_{b=1}^{n} W_b \right) + \left(\frac{1}{n} \sum_{1\leq i,j\leq n} f(Z_i, Z_j) \right) \left(\frac{1}{n} \sum_{b=1}^{n} W_b \right)^2 . \tag{69}$$

Last two terms converge to zero since by Lemma 18 $\left(\frac{1}{n}\sum_{b=1}^{n} W_b\right)$ converges to zero and by the Lemma 17 (with $g = 1$) $\frac{1}{n}\sum_{1\leq i,j\leq n} f(Z_i, Z_j)$ converges in law. The first term converges by the Lemma 17. □

A.2 Proof of the Theorem 1

Lemma 20. *Assume that the stationary process Z_t is τ-dependent with a coefficient $\tau(i) = i^{-6-\epsilon}$ for some $\epsilon > 0$. If core h is Lipschitz continuous, one-degenerate, bounded, and its h_2 component is a kernel then its normalized V statistic limiting distribution is proportional to its second component normalized V-statistic distribution. Shortly*

$$\lim_{n\to\infty} \varphi(nV(h), \binom{m}{2} nV(h_2)) = 0. \tag{70}$$

where φ denotes Prokhorov metric.

Proof. Lemma 7 shows how to write core h as a sum of its components h_i ,

$$nV(h) = nV(h_m) + \binom{m}{1} nV(h_{m-1}) + ... + \binom{m}{m-2} nV(h_2) + \binom{m}{m-1} nV(h_1) + h_0. \tag{71}$$

By Lemma 5 all components of h are bounded and Lipschitz continuous. Since h is one-degenerate, $h_0 = 0$ and component $h_1(z)$ is equal to zero everywhere

$$h_1(z) = \mathcal{E}h(z, Z_2^*, ..., Z_m^*) = 0. \tag{72}$$

By Lemma 13, for $c \geq 3$, $nV(h_c)$ converges to zero in probability. Therefore the behaviour of $V(h)$ is determined by $\binom{m}{m-2} V(h_2) = \binom{m}{2} V(h_2)$. Convergence of $V(h_2)$ follows from the [18, Theorem 2.1]. □

Lemma 21. *Assume that the stationary process Z_t is τ-dependent with a coefficient $\tau(i) = i^{-6-\epsilon}$ for some $\epsilon > 0$. If core h is Lipschitz continuous, one-degenerate, bounded, and its h_2 component is a kernel then its normalized and bootstrapped V statistic limiting distribution is same its second component normalized and bootstrapped V-statistic distribution. Shortly*

$$\lim_{n\to\infty} \varphi(nV_b(h), V_b(h_2)) = 0. \tag{73}$$

where φ denotes Prokhorov metric.

Proof. We show that the proposition holds for V_{b1} and then we prove that $\varphi(nV_{b2}(h), V_{b1}(h)) = 0$ converges to zero - the concept is to [18].

V_{b1} **convergence.** We write core h as a sum of components h_i (h_0, h_1 are equal to zero and therefore omitted). By Lemma 2

$$nV_{b1}(h) = \frac{1}{n^{m-1}} \sum_{i \in N^m} \Big[W_{i_1} W_{i_2} h_m(Z_{i_1}, ..., Z_{i_m}) + \tag{74}$$

$$\sum_{1 \leq j_1 < ... < j_{m-1} \leq m} W_{i_1} W_{i_2} h_{m-1}(Z_{i_{j_1}}, ..., Z_{i_{j_{m-1}}}) + ... + \sum_{1 \leq j_1 < j_2 \leq m} W_{i_1} W_{i_2} h_2(Z_{i_{j_1}}, Z_{i_{j_2}}) \Big]. \tag{75}$$

Consider a sum associated with h_2

$$\frac{1}{n^{m-1}} \sum_{i \in N^m} \sum_{1 \leq j_1 < j_2 \leq m} W_{i_1} W_{i_2} h_2(Z_{i_{j_1}}, Z_{i_{j_2}}). \tag{76}$$

Fix j_1, j_2. If $j_1 \neq 1$, $j_2 \neq 2$ then the sum

$$\frac{1}{n^{m-1}} \sum_{i \in N^m} W_{i_1} W_{i_2} h_2(Z_{i_{j_1}}, Z_{i_{j_2}}) \overset{L.6}{=\!=} \frac{1}{n^3} \sum_{i \in N^4} W_{i_1} W_{i_2} h_2(Z_{i_3}, Z_{i_4}) = \tag{77}$$

$$\left(\frac{1}{n} \sum_{i \in N^2} h_2(Z_{i_1}, Z_{i_2}) \right) (\frac{1}{n} \sum_{i=1}^{n} W_i)^2 \xrightarrow{L.15} 0 \text{ in probability.} \tag{78}$$

If $j_1 = 1$ and $j_2 \neq 2$, then the sum

$$\frac{1}{n^{m-1}} \sum_{i \in N^m} W_{i_1} W_{i_2} h_2(Z_{i_{j_1}}, Z_{i_{j_2}}) \overset{L.6}{=\!=} \frac{1}{n^2} \sum_{i \in N^3} W_{i_1} W_{i_3} h_2(Z_{i_1}, Z_{i_3}) = \tag{79}$$

$$\left(\frac{1}{n} \sum_{i \in N^2} W_{i_1} h_2(Z_{i_1}, Z_{i_2}) \right) \left(\frac{1}{n} \sum_{i=1}^{n} W_i \right) \xrightarrow{L.14,16} 0 \text{ in probability.} \tag{80}$$

The similar reasoning holds for $j_i = 2$ and $j_2 > 2$. The sum associated with h_c for $c > 2$

$$\frac{1}{n^{m-1}} \sum_{i \in N^m} \sum_{1 \leq j_1 < ... < j_c \leq m} W_{i_1} W_{i_2} h_c(Z_{i_{j_1}}, ..., Z_{i_{j_c}}) \xrightarrow{L.12} 0 \text{ in probability.} \tag{81}$$

Therefore

$$\lim_{n\to\infty}\left(nV_b(h)-\sum_{i\in N^2}W_{i_1}W_{i_2}h_2(Z_{i_1},Z_{i_2})\right)\overset{P}{=}0. \tag{82}$$

what proofs the proposition for V_{b1}.

V_{b1} **convergence.** To prove that V_{b2} converges to the same distribution as V_{b1} we investigate the difference

$$V_{b1}-V_{b2}=\frac{1}{n^{m-1}}\sum_{i\in N^m}W_{i_1}W_{i_2}h(Z_{i_1},...,Z_{i_m})-\frac{1}{n^{m-1}}\sum_{i\in N^m}\tilde{W}_{i_1}\tilde{W}_{i_2}h(Z_{i_1},...,Z_{i_m})= \tag{83}$$

$$\frac{1}{n^{m-1}}\sum_{i\in N^m}W_{i_1}W_{i_2}h(\cdot)-\frac{1}{n^{m-1}}\sum_{i\in N^m}(W_{i_1}-\frac{1}{n}\sum_{j=1}^n W_j)(W_{i_2}-\frac{1}{n}\sum_{j=1}^n W_j)h(\cdot)= \tag{84}$$

$$-\left(\frac{2}{n}\sum_{j=1}^n W_j\right)\left(\frac{1}{n^{m-1}}\sum_{i\in N^m}W_{i_1}h(\cdot)\right)+\left(\frac{1}{n^{m-1}}\sum_{i\in N^m}h(\cdot)\right)\left(\frac{1}{n}\sum_{j=1}^n W_j\right)^2. \tag{85}$$

The second term

$$\left(\frac{1}{n^{m-1}}\sum_{i\in N^m}h(Z_{i_1},...,Z_{i_m})\right)\left(\frac{1}{n}\sum_{j=1}^n W_j\right)^2\xrightarrow{L.15}0 \text{ in probability.} \tag{86}$$

Therefore we only need to show that the first term converges to zero

$$\left(\frac{2}{n}\sum_{j=1}^n W_j\right)\left(\frac{1}{n^{m-1}}\sum_{i\in N^m}W_{i_1}h(Z_{i_1},...,Z_{i_m})\right). \tag{87}$$

Since $\frac{2}{n}\sum_{j=1}^n W_j$ converges in probability to zero (by Lemma 14) we only nee to show that $\frac{1}{n^{m-1}}\sum_{i\in N^m}W_{i_1}h(Z_{i_1},...,Z_{i_m})$ converges. Using decomposition from Lemma 2 we write

$$\frac{1}{n^{m-1}}\sum_{i\in N^m}W_{i_1}h(Z_{i_1},...,Z_{i_m})=\frac{1}{n^{m-1}}\sum_{i\in N^m}\Big[W_{i_1}h_m(Z_{i_1},...,Z_{i_m})+ \tag{88}$$

$$\sum_{1\le j_1<...<j_{m-1}\le m}W_{i_1}h_{m-1}(Z_{i_{j_1}},...,Z_{i_{j_{m-1}}})+...+\sum_{1\le j_1<j_2\le m}W_{i_1}h_2(Z_{i_{j_1}},Z_{i_{j_2}})\Big]. \tag{89}$$

Term associated with h_2 can be written as

$$\frac{1}{n^{m-1}}\sum_{i\in N^m}\sum_{1\le j_1<j_2\le m}W_{i_1}h_2(Z_{i_{j_1}},Z_{i_{j_2}})= \tag{90}$$

$$=\begin{cases} n^{-1}\sum_{i\in N^2}W_{i_1}h_2(Z_{i_1},Z_{i_2}): j_1=1 \text{ or } j_1=2\\ \left(n^{-1}\sum_{i\in N^2}h_2(Z_{i_1},Z_{i_2})\right)\left(\frac{1}{n}\sum_{j=1}^n W_j\right): \text{ otherwise .}\end{cases} \tag{91}$$

In the first case ($j_1=1$ or $j_1=2$) Lemma 16 assures convergence. In the second case we use Lemma 15 to show convergence to zero. Other terms with h_c for $c>2$

$$\frac{1}{n^{m-1}}\sum_{i\in N^m}\sum_{1\le j_1<...<j_c\le m}W_{i_1}h_{m-1}(Z_{i_{j_1}},...,Z_{i_{j_c}})\xrightarrow{L.12}0 \text{ in probability.} \tag{92}$$

□

A.3 Proof of the Proposition 3

Lemma 22. *$nV_{b2}(h)$ converges to some non-zero random variable with finite variance.*

Proof. Using decomposition from the Lemma 2 we write core h as a sum of components h_c and h_0

$$nV_{b2}(h) = \frac{1}{n^{m-1}} \sum_{i\in N^m} \Big[h_0 \tilde{W}_{i_1}\tilde{W}_{i_2} + \sum_{1\le j\le m} \tilde{W}_{i_1}\tilde{W}_{i_2} h_1(Z_{i_j}) \tag{93}$$

$$\sum_{1\le j_1<j_2\le m} \tilde{W}_{i_1}\tilde{W}_{i_2} h_2(Z_{i_{j_1}}, Z_{i_{j_2}}) + ... + \tilde{W}_{i_1}\tilde{W}_{i_2} h_m(Z_{i_1}, ..., Z_{i_m}) \Big]. \tag{94}$$

We examine terms of the aboves sum starting form the one with h_0 - it is equal to zero

$$\frac{1}{n^{m-1}} \sum_{i\in N^m} h_0 \tilde{W}_{i_1}\tilde{W}_{i_2} \stackrel{L.6}{=\!=} \frac{1}{n} h_0 \sum_{i\in N^2} \tilde{W}_{i_1}\tilde{W}_{i_2} = \frac{1}{n} h_0 \left(\sum_{i=1} \tilde{W}_i \right)^2 \stackrel{L.18}{=\!=} 0. \tag{95}$$

Term with h_1 is zero as well, to see that fix j and consider

$$T_j = \frac{1}{n^{m-1}} \sum_{i\in N^m} \tilde{W}_{i_1}\tilde{W}_{i_2} h_1(Z_{i_j}). \tag{96}$$

If $j = 1$ then

$$T_1 \stackrel{L.6}{=\!=} \frac{1}{n} \sum_{i\in N^2} \tilde{W}_{i_1}\tilde{W}_{i_2} h_1(Z_{i_1}) = \frac{1}{n} \left(\sum_{i=1}^{n} \tilde{W}_i h_1(Z_i) \right) \left(\sum_{i=1} \tilde{W}_i \right) \stackrel{L.18}{=\!=} 0. \tag{97}$$

If $j = 2$ the same reasoning holds and if $j > 2$

$$T_j \stackrel{L.6}{=\!=} \frac{1}{n^2} \sum_{i\in N^3} \tilde{W}_{i_1}\tilde{W}_{i_2} h_1(Z_{i_3}) = \frac{1}{n} \left(\sum_{i=1}^{n} h_1(Z_i) \right) \left(\sum_{i=1} \tilde{W}_i \right)^2 \stackrel{L.18}{=\!=} 0. \tag{98}$$

Term containing h_2

$$T_{j_1,j_2} = \frac{1}{n^{m-1}} \sum_{i\in N^m} \tilde{W}_{i_1}\tilde{W}_{i_2} h_2(Z_{i_{j_1}}, Z_{i_{j_2}}) \tag{99}$$

is not zero. In the Lemma 19 we show that for $j_1 = 1$ and $j_2 = 2$ it converges to some non-zero variable. For $j_1 = 1$ and $j_2 > 2$ we have

$$T_{1,j_2} \stackrel{L.6}{=\!=} \frac{1}{n^2} \sum_{i\in N^3} \tilde{W}_{i_1}\tilde{W}_{i_2} h_2(Z_{i_1}, Z_{i_{j_2}}) = \frac{1}{n^2} \left(\sum_{i\in N^2} \tilde{W}_{i_1} h_2(Z_{i_1}, Z_{i_2}) \right) \left(\sum_{i=1} \tilde{W}_i \right) \stackrel{L.18}{=\!=} 0. \tag{100}$$

Exactly the same argument works for $T_{j_2,1}$. If both $j_1 \neq 1$ and $j_2 \neq 2$ then

$$T_{j_1,j_2} \stackrel{L.6}{=\!=} \frac{1}{n^3} \sum_{i\in N^4} \tilde{W}_{i_1}\tilde{W}_{i_2} h_2(Z_{i_{j_1}}, Z_{i_{j_2}}) = \frac{1}{n^3} \left(\sum_{i\in N^2} h_2(Z_{i_{j_1}}, Z_{i_{j_2}}) \right) \left(\sum_{i=1} \tilde{W}_i \right)^2 \stackrel{L.18}{=\!=} 0. \tag{101}$$

Terms containing h_c for $c > 2$

$$\frac{1}{n^{m-1}} \sum_{i\in N^m} \sum_{1\le j_1<...<j_c\le m} \tilde{W}_{i_1}\tilde{W}_{i_2} h_{m-1}(Z_{i_{j_1}}, ..., Z_{i_{j_c}}) \xrightarrow{L.12} 0 \tag{102}$$

converge to zero in probability. □

Lemma 23. *V_{b1} converges to zero in probability.*

Proof. The expected value and variance of V_{b1} converge to 0, therefore V_{b1} converges to zero in probability. Indeed for an expected value we have

$$\mathcal{E}V_{b1} = \frac{1}{n^m} \sum_{i\in N^m} \mathcal{E}W_{i_1}W_{i_2}\mathcal{E}h(Z_{i_1}, ..., Z_{i_m}) = \frac{1}{n^m} \sum_{i\in N^m} e^{|i_2-i_1|/ln} \mathcal{E}h(\cdot) \le \tag{103}$$

$$\frac{1}{n^m} \sum_{i\in N^m} e^{|i_2-i_1|/ln} \parallel h \parallel_\infty = \parallel h \parallel_\infty \frac{1}{n^2} \sum_{i\in N^2} e^{|i_2-i_1|/ln} \to 0. \tag{104}$$

Similar reasoning shows convergence of $\mathcal{E}V_{b1}^2$. □

A.4 Proof of Proposition 1

Proposition 5. *Let k be bounded and Lipschitz continuous, and let $\{X_t\}$ and $\{Y_t\}$ both be τ-dependent with coefficients $\tau(i) = o(\frac{1}{i^3})$, but independent of each other. Further, let $n_x = \rho_x n$ and $n_y = \rho_y n$ where $n = n_x + n_y$. Then, under the null hypothesis $P_x = P_y$, $\rho_x \rho_y n\widehat{MMD}_k$ and $\rho_x \rho_y n\widehat{MMD}_{k,b}$ converge to the same distribution as $n \to \infty$.*

Proof. Since $\widehat{\text{MMD}}_k$ is just the MMD between empirical measures using kernel k, it must be the same as the empirical MMD $\widehat{\text{MMD}}_{\tilde{k}}$ with centred kernel $\tilde{k}(x, x') = \langle k(\cdot, x) - \mathcal{E}k(\cdot, X), k(\cdot, x') - \mathcal{E}k(\cdot, X)\rangle_{\mathcal{H}_k}$ according to [22, Theorem 22]. Using the Mercer expansion, we can write

$$\begin{aligned}\rho_x \rho_y n\widehat{\text{MMD}}_k &= \rho_x \rho_y n \sum_{r=1}^{\infty} \lambda_r \left(\frac{1}{n_x} \sum_{i=1}^{n_x} \Phi_r(x_i) - \frac{1}{n_y} \sum_{j=1}^{n_y} \Phi_r(y_j) \right)^2 \\ &= \sum_{r=1}^{\infty} \lambda_r \left(\sqrt{\frac{\rho_y}{n_x}} \sum_{i=1}^{n_x} \Phi_r(x_i) - \sqrt{\frac{\rho_x}{n_y}} \sum_{j=1}^{n_y} \Phi_r(y_j) \right)^2,\end{aligned}$$

where $\{\lambda_r\}$ and $\{\Phi_r\}$ are the eigenvalues and the eigenfunctions of the integral operator $f \mapsto \int f(x)\tilde{k}(\cdot, x)dP_x(x)$ on $L_2(P_x)$. Similarly as in [18, Theorem 2.1], the above converges in distribution to $\sum_{r=1}^{\infty} \lambda_r Z_r^2$, where $\{Z_r\}$ are marginally standard normal, jointly normal and given by $Z_r = \sqrt{\rho_x} A_r - \sqrt{\rho_y} B_r$. $\{A_r\}$ and $\{B_r\}$ are in turn also marginally standard normal and jointly normal, with a dependence structure induced by that of $\{X_t\}$ and $\{Y_t\}$ respectively. This suggests individually bootstrapping each of the terms $\Phi_r(x_i)$ and $\Phi_r(y_j)$, giving rise to

$$\begin{aligned}\widehat{\text{MMD}}_{\tilde{k},b} &= \sum_{r=1}^{\infty} \lambda_r \left(\frac{1}{n_x} \sum_{i=1}^{n_x} \Phi_r(x_i)\tilde{W}_i^{(x)} - \frac{1}{n_y} \sum_{j=1}^{n_y} \Phi_r(y_j)\tilde{W}_j^{(y)} \right)^2 \\ &= \frac{1}{n_x^2} \sum_{i=1}^{n_x} \sum_{j=1}^{n_x} \tilde{W}_i^{(x)}\tilde{W}_j^{(x)}\tilde{k}(x_i, x_j) - \frac{1}{n_x^2} \sum_{i=1}^{n_y} \sum_{j=1}^{n_y} \tilde{W}_i^{(y)}\tilde{W}_j^{(y)}\tilde{k}(y_i, y_j) \\ &\quad - \frac{2}{n_x n_y} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \tilde{W}_i^{(x)}\tilde{W}_j^{(y)}\tilde{k}(x_i, y_j).\end{aligned}$$

Now, since $\tilde{k}$ is degenerate under the null distribution, the first two terms (after appropriate normalization) converge in distribution to $\rho_x \sum_{r=1}^{\infty} \lambda_r A_r^2$ and $\rho_y \sum_{r=1}^{\infty} \lambda_r B_r^2$ by [18, Theorem 3.1] as required. The last term follows the same reasoning - it suffices to check part (b) of [18, Theorem 3.1] (which is trivial as processes $\{X_t\}$ and $\{Y_t\}$ are assumed to be independent of each other) and apply the continuous mapping theorem to obtain convergence to $-2\sqrt{\rho_x\rho_y} \sum_{r=1}^{\infty} \lambda_r A_r B_r$ implying that $\widehat{\text{MMD}}_{\tilde{k},b}$ has the same limiting distribution as $\widehat{\text{MMD}}_k$. While we cannot compute $\tilde{k}$ as it depends on the underlying probability measure P_x. However, it is readily checked that due to the empirical centering of processes $\{\tilde{W}_t^{(x)}\}$ and $\{\tilde{W}_t^{(y)}\}$, $\widehat{\text{MMD}}_{\tilde{k},b} = \widehat{\text{MMD}}_{k,b}$, which proves the claim. Note that the result fails to be valid for non-empirically centred wild bootstrap processes. □

A.5 Lag-HSIC

Propositions 2 and 3 also allow us to construct a test of time series independence that is similar to one designed by [4]. Here, we will be testing against a broader null hypothesis: for all times t, t', X_t and $Y_{t'}$ are independent.

Since the time series $Z_t = (X_t, Y_t)$ is stationary, it suffices to check if there exists a dependency between X_t and Y_{t+k} for $-f_n \leq k \leq f_n$ where $\{f_n\}$ is an increasing sequence of positive numbers such that $f_n = o(n)$, but $\lim_{n\to\infty} f_n = \infty$. With increasing sample size n, we cover a wider range of lags. Since each lag corresponds to an individual hypothesis, we will require a multiple hypothesis testing correction to attain a desired test level α. We therefore define a sequence of quantiles given by $q_{\alpha,n} = 1 - \frac{\alpha}{2f_n+1}$. The shifted time series will be denoted $Z_t^k = (X_t, Y_{t+k})$. Further, let $S_{k,n} = nV(h, Z^k)$ be value of the HSIC statistic calculated on the shifted process Z_t^k. Finally let F_n and F denote respectively the finite-sample and the limiting distribution under the null hypothesis of the $S_{0,n} = nV(h,Z)$ (or, equivalently, of any $S_{k,n}$ since the null hypothesis holds).

Let us assume that we have computed the empirical $q_{\alpha,n}$-quantile of the null distribution, denoted by $t^*_{q_{\alpha,n}}$. The null hypothesis will then be rejected in the event

$$\mathcal{A}_{F,n} = \left\{ \max_{-f_n \le k \le f_n} S_{k,n} > t^*_{q_{\alpha,n}} \right\}.$$

Denote $\mathcal{A}_F = \lim_{n\to\infty} \mathcal{A}_{F,n}$. Our Theorem 1 implies that both $t^*_{q_{\alpha,n}}$ and $F_n^{-1}(q_{\alpha,n})$ converge to $F^{-1}(q_{\alpha,n})$. We will assume that $|q_{\alpha,n} - F_n(t^*_{q_{\alpha,n}})| = o(\frac{1}{n^r})$, for some $r > 0$, which follows from the Barry-Esseen-type argument for degenerate V-statistics on a τ-mixing process, but the proof of such result and the specific convergence rates are beyond the scope of this work. Then, by continuity and sub-additivity of probability, the asymptotic Type I error is given by

$$P_{H_0}(\mathcal{A}_F) = \lim_{n\to\infty} P_{H_0}(\mathcal{A}_{F,n}) \le \lim_{n\to\infty} \sum_{-f_n\le k\le f_n} P_{H_0}(S_{k,n} > t^*_{q_{\alpha,n}}) = \tag{105}$$

$$\lim_{n\to\infty} (2f_n+1)\left(1 - F_n(t^*_{q_{\alpha,n}})\right) \le \lim_{n\to\infty}(2f_n+1)\left(1 - (1 - \frac{\alpha}{2f_n+1}) + \frac{C}{n^r}\right) = \alpha, \tag{106}$$

as long as $f_n = o(n^r)$.

On the other hand if null hypothesis does not hold, there exists k for which $n^{-1}S_{k,n}$ converges to some positive constant. Unfortunately if the alternative hypothesis holds one can not approximate distribution F. We can however show that for any cumulative distribution function G, such that if $X \sim G$ and $\mathcal{E}X^2 < \infty$ the following holds

$$P(\mathcal{A}_G) = \lim_{n\to\infty} P(\mathcal{A}_{G,n}) \le \lim_{n\to\infty} P(S_{k,n} > G^{-1}(q_{\alpha,n})) = \lim_{n\to\infty} P(n^{-1}S_{k,n} > n^{-1}G^{-1}(q_{\alpha,n})) = 1.$$

The last equality follows from Lemma 24, which shows that $n^{-1}G^{-1}(q_{\alpha,n})$ converges to zero, and from the fact that $\frac{1}{n}S_{k,n}$ converges to some positive constant.

Lemma 24. *If $X \sim G$ is a random variable such that $\mathcal{E}X^2 < \infty$, $q_{\alpha,n} = 1 - \frac{\alpha}{2f_n+1}$ and $f_n = o(n)$ then $n^{-1}G^{-1}(q_{\alpha,n}) \to 0$.*

Proof. First observe that by Markov inequality $P(X \ge t) \le \frac{\mathcal{E}X^2}{t}$ and therefore $G(t) > g(t) = 1 - \frac{\mathcal{E}X^2}{t}$. Therefore, on the interval $(\mathcal{E}X, 1)$, $G^{-1}(x) < g^{-1}(x) = \frac{\mathcal{E}X^2}{1-x}$. As a result

$$n^{-1}G^{-1}(q_{\alpha,n}) \le n^{-1}g^{-1}(q_{\alpha,n}) = n^{-1}\frac{\mathcal{E}X^2}{1-(1-\frac{\alpha}{2f_n+1})} = \frac{(2f_n+1)\mathcal{E}X^2}{\alpha n} \overset{n\to\infty}{\longrightarrow} 0. \tag{107}$$

□

These observation result in the following test. We calculate an approximation of the cumulative distribution function F using $nV_{b2}(h)$. Under the null hypothesis, the Type I error is controlled by the bound from the equation (105). If the alternative holds, cumulative distribution function obtained from sampling $nV_{b2}(h)$ is G (possibly different from F). By proposition 3, $\int x^2 G(dx) < \infty$, and thus the control of the Type II error follows from the equation 107.