Supplement to "Does the Markov Decision Process Fit the Data: Testing for the Markov Property in Sequential Decision Making"

Chengchun Shi [1] **Runzhe Wan** [2] **Rui Song** [2] **Wenbin Lu** [2] **Ling Leng** [3]

Computing a single test in Section 5.2 takes roughly 10 minutes. Computing a single test in Section 5.3 takes roughly 30 seconds. The runtime can be further reduced with parallelization.

A. Additional details regarding our test

A.1. The covariance estimator $\widehat{\Sigma}^{(q)}$

For any $\ell = 1, \cdots, \mathbb{L}$, $j \in \mathcal{I}^{(\ell)}$ and $0 < t < T - q$, define vectors $\lambda_{R,q,j,t}, \lambda_{I,q,j,t} \in \mathbb{R}^{\mathbb{B}}$ such that the b-th element of $\lambda_{R,q,j,t}, \lambda_{I,q,j,t}$ correspond to the real and imaginary part of

$$\{\exp(i\mu^\top S_{j,t+q+1}) - \widehat{\varphi}^{(-\ell)}(\mu|X_{j,t+q})\}\{\exp(i\nu^\top X_{j,t-1}) - \widehat{\psi}^{(-\ell)}(\nu|X_{j,t})\},$$

respectively. The matrix $\widehat{\Sigma}^{(q)}$ is defined by

$$\sum_{\ell=1}^{\mathbb{L}} \sum_{j \in \mathcal{I}^{(\ell)}} \sum_{t=1}^{T-q-1} \frac{(\lambda_{R,q,j,t}^\top, \lambda_{I,q,j,t}^\top)^\top (\lambda_{R,q,j,t}^\top, \lambda_{I,q,j,t}^\top)}{n(T-q-1)}. \tag{A.1}$$

A.2. Validity of our test without the stationary assumption

When (C2) is violated, the relation $\psi_1 = \psi_2 = \cdots = \psi_{T-1}$ might no longer hold. However, under (C1), (C3) and H_0, the marginal distribution function of X_t can be well-approximated by some $\mathbb{F}$ on average. As a result, ψ_t's can be well-approximated by some ψ^* on average. Let $\mathbb{F}_t$ denote the distribution function of X_t. As long as the prediction error satisfies

$$\max_{1 \le b \le B} \frac{1}{T} \sum_{t=1}^{T} \int_x |\widehat{\psi}^{(-\ell)}(\nu_b|x) - \psi_t(\nu_b|x)|^2 \mathbb{F}_t(dx) = O_p((nT)^{-c_0}),$$

for some $c_0 > 1/2$, our test remains valid.

B. More on the OhioT1DM dataset

B.1. Detailed definitions of actions and rewards

We define $A_{i,t}$ as follows:

$$A_{i,t} = \begin{cases} 0, & \text{In}_{i,t} = 0; \\ m, & 4m - 4 < \text{In}_{i,t} \le 4m \quad (m = 1, 2, 3); \\ 4, & \text{In}_{i,t} > 12. \end{cases}$$

The Index of Glycemic Control is chosen as the immediate reward $R_{i,t}$, defined by

[1]Department of Statistics, London School of Economics and Political Science, UK [2]Department of Statistics, North Carolina State University, USA [3]Amazon, USA. Correspondence to: Chengchun Shi <c.shi7@lse.ac.uk>.

Proceedings of the 37^{th} International Conference on Machine Learning, Online, PMLR 119, 2020.

$$R_{i,t} = \begin{cases} -\frac{1}{30}(80 - \mathrm{G}_{i,t+1})^2, & \mathrm{G}_{i,t+1} < 80; \\ 0, & 80 \le \mathrm{G}_{i,t+1} \le 140; \\ -\frac{1}{30}(\mathrm{G}_{i,t+1} - 140)^{1.35}, & 140 \le \mathrm{G}_{i,t+1}. \end{cases}$$

B.2. Detailed procedure for value evaluation in simulations

In Section 5.2, we compare the policies learned with the selected order $\widehat{\kappa}_0$ and fixed orders $k \in \{1, \cdots, 10\}$. Below, we provide more details on computing the value $V^{(l)}(k)$.

1. In the l-th simulation, generate n trajectories $\{(S_{j,t}, A_{j,t})\}_{1\le j\le n, 0\le t\le 1344}$, and apply Algorithm 2 with $\alpha = 0.01$ and $K = 10$ to estimate an order $\widehat{\kappa}_0^{(l)}$. Also generate 100 trajectories of length 10 with the model described in Section 5.2, denoted by $\{(S^e_{j,t}, A^e_{j,t})\}_{1\le j\le 100, 0\le t<10}$.
2. For $k = 1, \ldots, 10$, apply FQI (see below) to the concatenated data $\{(S_{j,t}(k), A_{j,t}(k), R_{j,t}(k))\}_{1\le j\le n, 0\le t\le 1344-k}$ to learn an optimal policy $\widehat{\pi}^{(l)}(k)$.
3. For each initial trajectory $\{(S^e_{j,t}, A^e_{j,t})\}_{0\le t<10}$, generate the data $\{(S^e_{j,t}, A^e_{j,t}, R^e_{j,t})\}_{10\le t<60}$ following $\widehat{\pi}^{(l)}(k)$. Compute the value $V^{(l)}(k)$ by

$$V^{(l)}(k) = \frac{1}{100}\sum_{j=1}^{100}\sum_{t=10}^{50}\gamma^{t-10}R^e_{j,t},$$

with $\gamma = 0.9$.

Algorithm 1 Fitted-Q iteration

0: **Input:** Data $\{S_{j,t}, A_{j,t}, R_{j,t}, S_{j,t+1}\}_{j,t}$, function class $\mathcal{F}$, decay rate γ, action space $\mathcal{A}$
0: Randomly pick $Q_0 \in \mathcal{F}$
0: **For** $k = 1, \ldots, K$:
0: Update target values $Z_{j,t} = R_{j,t} + \gamma \max_{a\in\mathcal{A}} Q_{k-1}(S_{j,t+1}, a)$ for all (j, t);
0: Solve a regression problem to update the Q-function:
$Q_k = \arg\min_{Q\in\mathcal{F}} \frac{1}{n}\sum_{i=1}^{n}\{Q(S_{j,t}, A_{j,t}) - Z_{j,t}\}^2$
0: **Output:** The estimated optimal policy $\widehat{\pi}(\cdot) = \arg\max_{a\in\mathcal{A}} Q_K(\cdot, a)$

In our experiment,we use random forests to estimate the Q function during each iteration. The number of trees are set as 100 and the other hyperparameters are selected by 5-fold cross-validation. The decay rate γ is set to 0.9.

B.3. Detailed procedure for value evaluation in real data analysis

In Section 5.2, we compare policies learned by assuming the data follows a k-th order MDP for $k \in \{1, \cdots, 10\}$. The policies are estimated by FQI. To evaluate the values of these policies based on the real dataset, we apply the Fitted-Q evaluation (FQE) algorithm. Similar to FQI, it is an iterative algorithm based on the Bellman equation. We recap the steps below.

Algorithm 2 Fitted-Q evaluation

0: **Input:** Data $\{S_{j,t}, A_{j,t}, R_{j,t}, S_{j,t+1}\}_{j,t}$, policy π, function class $\mathcal{F}$, decay rate γ
0: Randomly pick $Q_0 \in \mathcal{F}$
0: **For** $k = 1, \ldots, K$
0: Update target values $Z_{j,t} = R_{j,t} + \gamma Q_{k-1}(S_{j,t+1}, \pi(S_{j,t+1}))$ for all (j, t);
0: Solve a regression problem to update the Q-function:
$Q_k = \arg\min_{Q\in\mathcal{F}} \frac{1}{n}\sum_{i=1}^{n}\{Q(S_{j,t}, A_{j,t}) - Z_{j,t}\}^2$
0: **Output:** The estimated value $\widehat{V}(\cdot) = Q_K(\cdot, \pi(\cdot))$

Denote the trajectories for the six patients in the OhioT1DM dataset by $\{(S_{i,t}, A_{i,t})\}_{1\le i\le 6, 1\le t\le 1100}$, and let the index set $\mathcal{I} = \{1,2,3,4,5,6\}$. We now describe the evaluation procedure in more details:

1. In $l = 1,\dots,20$, divide $\mathcal{I}$ into a training set $\mathcal{D}_1^{(l)}$ and an validation set $\mathcal{D}_2^{(l)} = (\mathcal{D}_1^{(l)})^c$ with $|\mathcal{D}_1^{(l)}| = |\mathcal{D}_2^{(l)}| = 3$.
2. For each $l \in \{1,\dots,20\}$, $k \in \{1,\dots,10\}$, apply FQI to the data $\{(S_{j,t}(k), A_{j,t}(k), R_{j,t}(k))\}_{j\in\mathcal{D}_1^{(l)}, 0\le t\le 1100-k+1}$ to learn an optimal policy $\widehat{\pi}^{(l)}(k)$.
3. For each $l \in \{1,\dots,20\}$, $k \in \{1,\dots,10\}$, apply FQE to the data $\{(S_{j,t}(k), A_{j,t}(k), R_{j,t}(k))\}_{j\in\mathcal{D}_2^{(l)}, 0\le t\le 1100-k+1}$ to estimate the state-value function of $\widehat{\pi}^{(l)}(k)$, denoted by $\widehat{V}_k^{(l)}(\cdot)$. Generate 100 trajectories of length 10 according to the simulation model in Section 5.2. Denote them by $\{(S_{j,t}^e, A_{j,t}^e)\}_{1\le j\le 100, 0\le t<10}$. Calculate the value under $\widehat{\pi}^{(l)}(k)$ by
$$V^{(l)}(k) = \frac{1}{100}\sum_{j=1}^{100} \widehat{V}_k^{(l)}(S_{j,(10-k)}^e(k)).$$
4. Average over the 20 splits to compute the average value for each k by $V(k) = \sum_{l=1}^{20} V^{(l)}(k)/20$.

For both FQI and FQE, we use random forests to estimate the regression function. The number of trees are set to 75 and the other hyperparameters are selected by 5-fold cross-validation. We set $\gamma = 0.9$ in our experiments.

B.4. More on the estimated values in real data analysis

We report the value difference between the 4-order MDP model and models with other orders and their associated standard errors in the following table. The standard errors are calculated according to the sample variance estimator of values obtained from 20 different combinations. Using a one-sided Z-test,we find that the gain of the value under the 4-order MDP model is not significant compared to model with 5-th or 8-th order. However, it is significant compared to all other models under the 0.1 significance level.

Table 1. Value difference for the OhioT1DM dataset.

k	1	2	3	5	6	7	8	9	10
Value difference	-38.25	-4.96	-11.20	-3.66	-7.48	-11.13	-2.28	-12.51	-7.02
Standard error	4.65	3.17	6.89	2.88	2.7	4.33	1.86	3.05	2.54

C. Technical proofs

C.1. Proof of Lemma 1

Consider a policy $\pi = \{\pi_t\}_{t\ge 0} \in$ HR. Suppose there exists some $\{\pi_t^*\}_{t\ge 0}$ such that $\pi_t(\cdot|\bar{\boldsymbol{S}}_t) = \pi_t^*(\cdot|S_t)$ almost surely for any $t \ge 0$. We refer to such a policy π as a *Markov* policy. In addition, π is a *deterministic* policy if and only if $\pi_t(a|\bar{\boldsymbol{S}}_t) \in \{0,1\}$ almost surely for any $t \ge 0$ and $a \in \mathcal{A}$. Let MR denotes the set of Markov policies and SD denote the set of deterministic stationary policies, we have SD $\subseteq$ SR $\subseteq$ MR $\subseteq$ HR. In the following, we focus on proving
$$\sup_{\pi\in \text{HR}} V(\pi; s) = \sup_{\pi\in \text{SD}} V(\pi; s), \qquad \forall s \in \mathbb{S}.$$

Since SD $\subseteq$ SR, the assertion in Lemma 1 is thus satisfied.

We begin by providing an outline of the proof. Our proof is divided into three steps. In the first step, we show
$$\sup_{\pi\in \text{HR}} V(\pi; s) = \sup_{\pi\in \text{MR}} V(\pi; s), \qquad \forall s \in \mathbb{S}.$$

To prove this, we show in Section C.1.1 that for any such $\pi \in$ HR and any s, there exists a Markov policy $\pi^* = \{\pi_t^*\}_{t\ge 0}$ where each π_t^* depends on S_t only such that
$$\mathbb{P}^{\pi}(A_t = a, S_t \in \mathcal{S}|S_0 = s) = \mathbb{P}^{\pi^*}(A_t = a, S_t \in \mathcal{S}|S_0 = s), \tag{C.1}$$

for any $t \geq 0$, $a \in \mathcal{A}$, $\mathcal{S} \subseteq \mathbb{S}$ and $s \in \mathbb{S}$ where the probabilities $\mathbb{P}^{\pi}$ and $\mathbb{P}^{\pi^*}$ are taken by assuming the system dynamics follow π and π^*, respectively. Under MA, we have

$$\mathbb{E}^{\pi}(R_{0,t}|S_0 = s) = \mathbb{E}^{\pi}\{\mathbb{E}^{\pi}(R_{0,t}|A_t, S_t, S_0 = s)|S_0 = x\} = \mathbb{E}^{\pi}\{r(A_t, S_t)|S_0 = x\},$$

for some function r. This together with (C.1) yields that

$$\mathbb{E}^{\pi}(R_{0,t}|S_0 = s) = \mathbb{E}^{\pi^*}(Y_{0,t}|S_0 = s), \quad \forall t \geq 0,$$

and hence $V(\pi; s) = V(\pi^*; s)$. This completes the proof for the first step.

With a slight abuse of notation, for any $\pi \in \text{SD}$, we denote by $\pi(s)$ the action that the agent chooses according to π, given that the current state equals s. In the second step, we show for any bounded function $\nu(\cdot)$ on $\mathbb{S}$ that satisfies the optimal Bellman equation

$$\nu(s) = \sup_{\pi \in \text{SD}} \left\{ r(\pi(s), s) + \gamma \int_{s'} \nu(s')\mathcal{P}(ds'; \pi(s), s) \right\}, \quad \forall s \in \mathbb{S},$$

it satisfies

$$\nu(s) = \sup_{\pi^* \in \text{MR}} V(\pi^*; s), \quad \forall s \in \mathbb{S}. \tag{C.2}$$

The proof of (C.2) is given in Section C.1.2.

For any function ν, define the norm $\|\nu\|_\infty = \sup_{s\in\mathbb{S}} |\nu(s)|$. We have for any ν_1 and ν_2 that

$$\begin{aligned}
\sup_x \left| \sup_{\pi \in \text{SD}} \left\{ r(\pi(s), s) + \gamma \int_{s'} \nu_1(s')\mathcal{P}(ds'; \pi(s), s) \right\} - \sup_{\pi \in \text{SD}} \left\{ r(\pi(s), s) + \gamma \int_{s'} \nu_2(s')\mathcal{P}(ds'; \pi(s), s) \right\} \right| \\
\leq \gamma \sup_{\pi \in \text{SD}} \sup_{s \in \mathbb{S}} \left| \int_{s'} \nu_1(s')\mathcal{P}(ds'; \pi(s), s) - \int_{s'} \nu_2(s')\mathcal{P}(ds'; \pi(s), s) \right| \\
\leq \gamma \sup_{\pi \in \text{SD}} \sup_{s \in \mathbb{S}} \left| \int_{s'} \|\nu_1 - \nu_2\|_\infty \mathcal{P}(ds'; \pi(s), s) \right| \leq \gamma \|\nu_1 - \nu_2\|_\infty.
\end{aligned}$$

By Banach's fix point theorem, there exists a unique value function ν_0 that satisfies the optimal Bellman equation. Combining this together with the results obtained in the first two steps, we obtain that ν_0 satisfies $\nu_0(s) = \sup_{\pi \in \text{HR}} V(\pi; s)$ for any $s \in \mathbb{S}$. The proof is thus completed if we can show there exists a deterministic stationary policy π^{**} that satisfies

$$\nu_0(s) = V(\pi^{**}; s), \quad \forall s \in \mathbb{S}. \tag{C.3}$$

We put the proof of (C.3) in Section C.1.3.

C.1.1. Proof of (C.1)

Apparently, (C.1) holds with $t = 0$. Suppose (C.1) holds for $t = k$. We now show (C.1) holds for $t = k + 1$. Under MA, we have

$$\begin{aligned}
& \mathbb{P}^{\pi}(S_{k+1} \in \mathcal{S}|S_0 = s) = \mathbb{E}^{\pi}\{\mathbb{P}^{\pi}(S_{t+1} \in \mathcal{S}|A_t, S_t, S_0 = s)|S_0 = x\} \\
= \quad & \mathbb{E}^{\pi}\{\mathcal{P}(\mathcal{S}; A_t, S_t)|S_0 = x\} = \mathbb{E}^{\pi^*}\{\mathcal{P}(\mathcal{S}; A_t, S_t)|S_0 = x\} = \mathbb{P}^{\pi^*}(S_{k+1} \in \mathcal{S}|S_0 = s) \triangleq \mathbb{G}_{k+1}(\mathcal{S}; s).
\end{aligned}$$

Set π^*_{k+1} to be the decision rule that satisfies

$$\mathbb{P}^{\pi}(A_{k+1} = a|S_{k+1}, S_0 = s) = \mathbb{P}^{\pi^*_{k+1}}(A_{k+1} = a|S_{k+1}), \quad \forall a \in \mathcal{A},$$

it follows that

$$\begin{aligned}
\mathbb{P}^{\pi}(A_{k+1} = a, S_{k+1} \in \mathcal{S}|S_0 = s) = \int_{s'} \mathbb{P}^{\pi}(A_{k+1} = a|S_{k+1} = s', S_0 = s)\mathbb{G}_{k+1}(ds'; s) \\
= \int_{s'} \mathbb{P}^{\pi^*}(A_{k+1} = a|S_{k+1} = s', S_0 = s)\mathbb{G}_{k+1}(ds'; s) = \mathbb{P}^{\pi^*}(A_{k+1} = a, S_{k+1} \in \mathcal{S}|S_0 = s).
\end{aligned}$$

Thus, (C.1) holds for $t = k + 1$ as well. The proof is hence completed.

C.1.2. PROOF OF (C.2)

We first show for any bounded function ν that satisfies

$$\nu(s) \geq \sup_{\pi \in \text{SD}} \left\{ r(\pi(s), s) + \gamma \int_{s'} \nu(s') \mathcal{P}(ds'; \pi(s), s) \right\}, \qquad \forall s \in \mathbb{S}, \tag{C.4}$$

we have

$$\nu(s) \geq \sup_{\pi^* \in \text{MR}} V(\pi^*; s), \qquad \forall s \in \mathbb{S}. \tag{C.5}$$

Then, we show for any bounded function ν that satisfies

$$\sup_{s \in \mathbb{S}} \left[\nu(s) - \sup_{\pi \in \text{SD}} \left\{ r(\pi(s), s) + \gamma \int_{s'} \nu(s') \mathcal{P}(ds'; \pi(s), s) \right\} \right] \leq 0,$$

we have

$$\nu(s) \leq \sup_{\pi^* \in \text{MR}} V(\pi^*; s), \qquad \forall s \in \mathbb{S}. \tag{C.6}$$

The proof is hence completed.

Proof of (C.5)*:* Consider an arbitrary deterministic Markov policy $\pi^* = \{\pi_t^*\}_{t \geq 0}$. With a slight abuse of notation, we denote by $\pi_t^*(s)$ the action that the agent chooses following π_t^*, given that the current state equals s. It follows from (C.4) that

$$\nu(s) \geq r(\pi_0^*(s), s) + \gamma \int_{s'} \nu(s') \mathcal{P}(ds'; \pi_0^*(s), s), \qquad \forall s \in \mathbb{S}.$$

By iteratively applying (C.4), we have

$$\nu(s) \geq r(\pi_0^*(s), s) + \sum_{k=1}^{K} \gamma^k \mathbb{E}^{\pi^*} \{r(A_k, X_k) | S_0 = x\} + \gamma^{K+1} \mathbb{E}^{\pi^*} \{\nu(X_{K+1}) | S_0 = x\}, \qquad \forall s \in \mathbb{S}.$$

Since ν is bounded, the last term on the right-hand-side (RHS) converges to zero uniformly in x, as $K \to \infty$. Let $K \to \infty$, we obtain $\nu(s) \geq V(\pi^*; s)$, for any $s \in \mathbb{S}$ and any deterministic Markov policy π^*. Using Lemma 4.3.1 of Puterman (1994), we can similarly show $\nu(s) \geq V(\pi^*; s)$ for any $s \in \mathbb{S}$ and $\pi^* \in$ MR. This completes the proof of (C.5).

Proof of (C.6)*:* By definition, we have

$$\inf_{\pi \in \text{SD}} \sup_{s \in \mathbb{S}} \left[\nu(s) - \left\{ r(\pi(s), s) + \gamma \int_{s'} \nu(s') \mathcal{P}(ds'; \pi(s), s) \right\} \right] \leq 0.$$

Thus, for any $\epsilon > 0$, there exists some $\pi_0 \in$ SD that satisfies

$$\sup_{s \in \mathbb{S}} \left[\nu(s) - \left\{ r(\pi_0(s), s) + \gamma \int_{s'} \nu(s') \mathcal{P}(ds'; \pi_0(s), s) \right\} \right] \leq \epsilon. \tag{C.7}$$

Consider the following bounded linear operator $\mathcal{L}_0$,

$$\mathcal{L}_0 \nu(s) = \int_{s'} \nu(s') \mathcal{P}(ds'; \pi_0(s), s),$$

defined on the space of bounded functions. Let $\mathcal{I}_0$ denote the identity operator. Since $\gamma < 1$, the operator $\mathcal{I}_0 - \gamma \mathcal{L}_0$ is invertible and its inverse equals $\sum_{k=0}^{+\infty} \gamma^k \mathcal{L}_0^k$. It follows from (C.7) that

$$\nu(s) \leq \sum_{k=0}^{+\infty} \gamma^k \mathcal{L}_0^k \{r(\pi_0(s), s) + \epsilon\}, \qquad \forall s \in \mathbb{S}.$$

Since $V(\pi_0; s) = \sum_{k=0}^{+\infty} \gamma^k \mathcal{L}_0^k r(\pi_0(s), s)$ and $\sum_{k=0}^{+\infty} \gamma^k \mathcal{L}_0^k \epsilon \leq \epsilon / (1 - \gamma)$, we obtain

$$\nu(s) \leq V(\pi_0; s) + \frac{\epsilon}{1 - \gamma}.$$

Let $\epsilon \to 0$, we obtain $\nu(s) \leq \sup_{\pi^* \in \text{MR}} V(\pi^*; s)$ for any x. The proof is hence completed.

C.1.3. PROOF OF (C.3)

Since $\nu_0(\cdot)$ satisfies the optimal Bellman equation, we have

$$\nu_0(s) = \arg\max_{\pi \in \text{SD}} \left\{ r(\pi(s), s) + \gamma \int_{s'} \nu_0(s') \mathcal{P}(ds'; \pi(s), s) \right\}.$$

Let $\mathcal{A}_s$ be the available set of actions at a given state s. As a result, we have

$$\nu_0(s) = \arg\max_{a \in \mathcal{A}_s} \left\{ r(a, s) + \gamma \int_{s'} \nu_0(s') \mathcal{P}(ds'; a, s) \right\}.$$

Since $\mathcal{A}$ is finite, so is $\mathcal{A}_s$. As a result, the above argmax is achievable. Let $\pi^{**}(s)$ be the action such that the above argmax is achieved, we have

$$\nu_0(s) = r(\pi^{**}(s), s) + \gamma \int_{s'} \nu_0(s') \mathcal{P}(ds'; \pi^{**}(s), s).$$

Similar to the proof of (C.2), we can show $\nu_0(s) = V(\pi^{**}; s)$, for all $s \in \mathbb{S}$. The proof is hence completed.

C.2. Proof of Lemma 2

Let $\widetilde{Z}_1, \widetilde{Z}_2$ be independent copies of Z_1, Z_2 such that $\widetilde{Z}_1 | Z_3 \stackrel{d}{=} Z_1 | Z_3$, $\widetilde{Z}_2 | Z_3 \stackrel{d}{=} Z_2 | Z_3$ and that $\widetilde{Z}_1 \perp\!\!\!\perp \widetilde{Z}_2 | \widetilde{Z}_3$. Consider any $\mu_1 \in \mathbb{R}^{q_1}, \mu_2 \in \mathbb{R}^{q_2}, \mu_3 \in \mathbb{R}^{q_3}$, we have

$$\begin{aligned} \mathbb{E}\exp(i\mu_1^\top \widetilde{Z}_1 + i\mu_2^\top \widetilde{Z}_2 + i\mu_3^\top Z_3) &= \mathbb{E}[\exp(i\mu_3^\top Z_3) \mathbb{E}\{\exp(i\mu_1^\top \widetilde{Z}_1 + i\mu_2^\top \widetilde{Z}_2) | Z_3\}] \\ &= \mathbb{E}[\exp(i\mu_3^\top Z_3) \mathbb{E}\{\exp(i\mu_1^\top \widetilde{Z}_1) | Z_3\} \mathbb{E}\{\exp(i\mu_2^\top \widetilde{Z}_2) | Z_3\}] \\ &= \mathbb{E}[\exp(i\mu_3^\top Z_3) \mathbb{E}\{\exp(i\mu_1^\top Z_1) | Z_3\} \mathbb{E}\{\exp(i\mu_2^\top Z_2) | Z_3\}]. \end{aligned} \tag{C.8}$$

Under the condition in Lemma 2, we have

$$\begin{aligned} \mathbb{E}[\exp(i\mu_3^\top Z_3) \mathbb{E}\{\exp(i\mu_1^\top Z_1) | Z_3\} \mathbb{E}\{\exp(i\mu_2^\top Z_2) | Z_3\}] &= \mathbb{E}[\exp(i\mu_3^\top Z_3) \mathbb{E}\{\exp(i\mu_1^\top Z_1 + i\mu_2^\top Z_2) | Z_3\}] \\ &= \mathbb{E}\exp(i\mu_1^\top Z_1 + i\mu_2^\top Z_2 + i\mu_3^\top Z_3). \end{aligned}$$

This together with (C.8) yields

$$\mathbb{E}\exp(i\mu_1^\top \widetilde{Z}_1 + i\mu_2^\top \widetilde{Z}_2 + i\mu_3^\top Z_3) = \mathbb{E}\exp(i\mu_1^\top Z_1 + i\mu_2^\top Z_2 + i\mu_3^\top Z_3).$$

As a result, $(\widetilde{Z}_1, \widetilde{Z}_2, Z_3)$ and (Z_1, Z_2, Z_3) have same characteristic functions. Therefore, we have $(\widetilde{Z}_1, \widetilde{Z}_2, Z_3) \stackrel{d}{=} (Z_1, Z_2, Z_3)$. By construction, we have $\widetilde{Z}_1 \perp\!\!\!\perp \widetilde{Z}_2 | Z_3$. It follows that $Z_1 \perp\!\!\!\perp Z_2 | Z_3$.

C.3. Proof of Theorem 3

We focus on proving Theorem 3 in the more challenging setting where $T \to \infty$. The number of trajectories n can be either bounded or growing to ∞. The case where T is bounded can be proven using similar arguments. We begin by providing an outline of the proof. For any q, μ, ν, define

$$\Gamma^*(q, \mu, \nu) = \frac{1}{n(T-q-1)} \sum_{j=1}^{n} \sum_{t=1}^{T-q-1} \{\exp(i\mu^\top S_{j,t+q+1}) - \varphi^*(\mu | X_{j,t+q})\} \{\exp(i\nu^\top X_{j,t-1}) - \psi^*(\nu | X_{j,t})\}.$$

Denote by Γ_R^* and Γ_I^* the real and imaginary part of Γ^*, respectively.

We break the proof into three steps. In the first step, we show

$$\max_{b \in \{1, \cdots, B\}} \max_{q \in \{0, \cdots, Q\}} \sqrt{n(T-q-1)} |\widehat{\Gamma}(q, \mu_b, \nu_b) - \Gamma^*(q, \mu_b, \nu_b)| = o_p(\log^{-1/2}(nT)). \tag{C.9}$$

Proof of (C.9) relies largely on Condition (C4) which requires $\widehat{\varphi}$ and $\widehat{\psi}$ to satisfy certain uniform convergence rates. This further implies that

$$\widehat{S} = S^* + o_p(\log^{-1/2}(nT)), \tag{C.10}$$

where

$$S^* = \max_{b\in\{1,\cdots,B\}} \max_{q\in\{0,\cdots,Q\}} \sqrt{n(T-q-1)}\max(|\Gamma_R^*(q,\mu_b,\nu_b)|, |\Gamma_I^*(q,\mu_b,\nu_b)|).$$

In the second step, we show for any $z \in \mathbb{R}$ and any sufficiently small $\varepsilon > 0$,

$$\mathbb{P}(S^* \le z) \ge \mathbb{P}(\|N(0,V_0)\|_\infty \le z - \varepsilon \log^{-1/2}(nT)) - o(1),$$
$$\mathbb{P}(S^* \le z) \le \mathbb{P}(\|N(0,V_0)\|_\infty \le z + \varepsilon \log^{-1/2}(nT)) + o(1),$$

where the matrix V_0 is defined in Step 2 of the proof. This together with (C.10) yields that

$$\mathbb{P}(\widehat{S} \le z) \ge \mathbb{P}(\|N(0,V_0)\|_\infty \le z - 2\varepsilon \log^{-1/2}(nT)) - o(1), \tag{C.11}$$
$$\mathbb{P}(\widehat{S} \le z) \le \mathbb{P}(\|N(0,V_0)\|_\infty \le z + 2\varepsilon \log^{-1/2}(nT)) + o(1). \tag{C.12}$$

The proposed Bootstrap algorithm repeatedly generate random variables from $\|N(0,\widehat{V})\|_\infty$ where the detailed form of $\widehat{V}$ is given in the third step of the proof. The critical values $\widehat{c}_\alpha$ is chosen to be the upper α-th quantile of $\|N(0,\widehat{V})\|_\infty$. In the third step, we show $\|V_0 - \widehat{V}\|_{\infty,\infty} = O((nT)^{-c^{**}})$ for some $c^{**} > 0$ with probability tending to 1, where $\|\cdot\|_{\infty,\infty}$ denotes the elementwise max-norm. Combining this upper bound with some arguments used in proving (C.11) and (C.12), we can show with probability tending to 1 that

$$\mathbb{P}(\widehat{S} \le z) \ge \mathbb{P}(\|N(0,\widehat{V})\|_\infty \le z - 2\varepsilon \log^{-1/2}(nT)|\widehat{V}) - o(1),$$
$$\mathbb{P}(\widehat{S} \le z) \le \mathbb{P}(\|N(0,\widehat{V})\|_\infty \le z + 2\varepsilon \log^{-1/2}(nT)|\widehat{V}) + o(1),$$

for any sufficiently small $\varepsilon > 0$ where $\mathbb{P}(\cdot|\widehat{V})$ denotes the conditional probability given $\widehat{V}$. Set $z = \widehat{c}_\alpha$. It follows from that

$$\mathbb{P}(\widehat{S} \le \widehat{c}_\alpha) \ge \mathbb{P}(\|N(0,\widehat{V})\|_\infty \le \widehat{c}_\alpha - 2\varepsilon \log^{-1/2}(nT)|\widehat{V}) - o(1), \tag{C.13}$$
$$\mathbb{P}(\widehat{S} \le \widehat{c}_\alpha) \le \mathbb{P}(\|N(0,\widehat{V})\|_\infty \le \widehat{c}_\alpha + 2\varepsilon \log^{-1/2}(nT)|\widehat{V}) + o(1), \tag{C.14}$$

with probability tending to 1. Under the given conditions in Theorem 3, the diagonal elements in V_0 are bounded away from zero. With probability tending to 1, the diagonal elements in $\widehat{V}$ is bounded away from zero as well. It follows from Theorem 1 of (Chernozhukov et al., 2017) that conditional on $\widehat{V}$,

$$\begin{aligned}\mathbb{P}(\|N(0,\widehat{V})\|_\infty \le \widehat{c}_\alpha + 2\varepsilon \log^{-1/2}(nT)|\widehat{V}) - \mathbb{P}(\|N(0,\widehat{V})\|_\infty \le \widehat{c}_\alpha - 2\varepsilon \log^{-1/2}(nT)|\widehat{V})\\ \le O(1)\varepsilon \log^{1/2}(BQ)\log^{-1/2}(nT),\end{aligned}$$

with probability tending to 1, where $O(1)$ denotes some positive constant that is independent of ε. Under the given conditions on B and Q, we obtain with probability tending to 1 that,

$$\mathbb{P}(\|N(0,\widehat{V})\|_\infty \le \widehat{c}_\alpha + 2\varepsilon \log^{-1/2}(nT)|\widehat{V}) - \mathbb{P}(\|N(0,\widehat{V})\|_\infty \le \widehat{c}_\alpha - 2\varepsilon \log^{-1/2}(nT)|\widehat{V}) \le C^*\varepsilon,$$

for some constant $C^* > 0$. This together with (C.13) and (C.14) yields

$$|\mathbb{P}(\widehat{S} \le \widehat{c}_\alpha) - \mathbb{P}(\|N(0,\widehat{V})\|_\infty \le \widehat{c}_\alpha|\widehat{V})| \le C^*\varepsilon + o(1),$$

with probability tending to 1. Notice that ε can be made arbitrarily small. The validity of our test thus follows.

In the following, we present our proof for each of the step. Suppose $\{\mu_b,\nu_b\}_{1\le b\le B}$ are fixed throughout the proof. Denote by $\widehat{\varphi}_R^{(\ell)}$, $\widehat{\varphi}_I^{(\ell)}$ the real and imaginary part of $\widehat{\varphi}^{(\ell)}$ respectively. Without loss of generality, we assume the absolute values of $\widehat{\varphi}_R^{(\ell)}$, $\widehat{\varphi}_I^{(\ell)}$ are uniformly bounded by 1.

C.3.1. Step 1

With some calculations, we can show that for any q,μ,ν,

$$\widehat{\Gamma}(q,\mu,\nu) = \Gamma^*(q,\mu,\nu) + R_1(q,\mu,\nu) + R_2(q,\mu,\nu) + R_3(q,\mu,\nu),$$

where the remainder terms R_1, R_2 and R_3 are given by

$$R_1(q,\mu,\nu) = \frac{1}{n(T-q-1)} \sum_{\ell=1}^{\mathbb{L}} \sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^{T-q-1} \{\varphi^*(\mu|X_{j,t+q}) - \widehat{\varphi}^{(-\ell)}(\mu|X_{j,t+q})\}\{\psi^*(\nu|X_{j,t}) - \widehat{\psi}^{(-\ell)}(\nu|X_{j,t})\},$$

$$R_2(q,\mu,\nu) = \frac{1}{n(T-q-1)} \sum_{\ell=1}^{\mathbb{L}} \sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^{T-q-1} \{\exp(i\mu^\top S_{j,t+q+1}) - \varphi^*(\mu|X_{j,t+q})\}\{\psi^*(\nu|X_{j,t}) - \widehat{\psi}^{(-\ell)}(\nu|X_{j,t})\},$$

$$R_3(q,\mu,\nu) = \frac{1}{n(T-q-1)} \sum_{\ell=1}^{\mathbb{L}} \sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^{T-q-1} \{\varphi^*(\mu|X_{j,t+q}) - \widehat{\varphi}^{(-\ell)}(\mu|X_{j,t+q})\}\{\exp(i\nu^\top X_{j,t-1}) - \psi^*(\nu|X_{j,t})\}.$$

It suffices to show

$$\max_{b\in\{1,\cdots,B\}} \max_{q\in\{0,\cdots,Q\}} \sqrt{n(T-q-1)}|R_m(q,\mu_b,\nu_b)| = o_p(\log^{-1/2}(nT)), \tag{C.15}$$

for $m=1,2,3$. In the following, we show (C.15) holds with $m=1,2$. Using similar arguments, one can show (C.15) holds with $m=3$.

Proof of (C.15) with $m=1$: Since $\mathbb{L}$ is fixed, it suffices to show

$$\max_{b\in\{1,\cdots,B\}} \max_{q\in\{0,\cdots,Q\}} \sqrt{n(T-q-1)}|R_{1,\ell}(q,\mu_b,\nu_b)| = o_p(\log^{-1/2}(nT)), \tag{C.16}$$

where $R_{1,\ell}(q,\mu_b,\nu_b)$ is defined by

$$\frac{1}{n(T-q-1)} \sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^{T-q-1} \{\varphi^*(\mu_b|X_{j,t+q}) - \widehat{\varphi}^{(-\ell)}(\mu_b|X_{j,t+q})\}\{\psi^*(\nu_b|X_{j,t}) - \widehat{\psi}^{(-\ell)}(\nu_b|X_{j,t})\}.$$

Similarly, let φ_R^* and φ_I^* denote the real and imaginary part of φ^*. We can rewrite $R_{1,\ell}(q,\mu_b,\nu_b)$ as $R_{1,\ell}^{(1)}(q,\mu_b,\nu_b) - R_{1,\ell}^{(2)}(q,\mu_b,\nu_b) + iR_{1,\ell}^{(3)}(q,\mu_b,\nu_b) + iR_{1,\ell}^{(4)}(q,\mu_b,\nu_b)$ where

$$R_{1,\ell}^{(1)}(q,\mu_b,\nu_b) = \frac{n^{-1}}{(T-q-1)} \sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^{T-q-1} \{\varphi_R^*(\mu_b|X_{j,t+q}) - \widehat{\varphi}_R^{(-\ell)}(\mu_b|X_{j,t+q})\}\{\psi_R^*(\nu_b|X_{j,t}) - \widehat{\psi}_R^{(-\ell)}(\nu_b|X_{j,t})\},$$

$$R_{1,\ell}^{(2)}(q,\mu_b,\nu_b) = \frac{n^{-1}}{(T-q-1)} \sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^{T-q-1} \{\varphi_I^*(\mu_b|X_{j,t+q}) - \widehat{\varphi}_I^{(-\ell)}(\mu_b|X_{j,t+q})\}\{\psi_I^*(\nu_b|X_{j,t}) - \widehat{\psi}_I^{(-\ell)}(\nu_b|X_{j,t})\},$$

$$R_{1,\ell}^{(3)}(q,\mu_b,\nu_b) = \frac{n^{-1}}{(T-q-1)} \sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^{T-q-1} \{\varphi_R^*(\mu_b|X_{j,t+q}) - \widehat{\varphi}_R^{(-\ell)}(\mu_b|X_{j,t+q})\}\{\psi_I^*(\nu_b|X_{j,t}) - \widehat{\psi}_I^{(-\ell)}(\nu_b|X_{j,t})\},$$

$$R_{1,\ell}^{(4)}(q,\mu_b,\nu_b) = \frac{n^{-1}}{(T-q-1)} \sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^{T-q-1} \{\varphi_R^*(\mu_b|X_{j,t+q}) - \widehat{\varphi}_R^{(-\ell)}(\mu_b|X_{j,t+q})\}\{\psi_I^*(\nu_b|X_{j,t}) - \widehat{\psi}_I^{(-\ell)}(\nu_b|X_{j,t})\}.$$

To prove (C.16), it suffices to show

$$\max_{b\in\{1,\cdots,B\}} \max_{q\in\{0,\cdots,Q\}} \sqrt{n(T-q-1)}|R_{1,\ell}^{(s)}(q,\mu_b,\nu_b)| = o_p(\log^{-1/2}(nT)), \tag{C.17}$$

for $s=1,2,3,4$. For brevity, we only show (C.17) holds with $s=1$.

By the Cauchy-Schwarz inequality, it suffices to show

$$\max_{b\in\{1,\cdots,B\}} \max_{q\in\{0,\cdots,Q\}} \frac{1}{\sqrt{n(T-q-1)}} \sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^{T} \{\varphi_R^*(\mu_b|X_{j,t}) - \widehat{\varphi}_R^{(-\ell)}(\mu_b|X_{j,t})\}^2 = o_p(\log^{-1/2}(nT)), \tag{C.18}$$

$$\max_{b\in\{1,\cdots,B\}} \max_{q\in\{0,\cdots,Q\}} \frac{1}{\sqrt{n(T-q-1)}} \sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^{T} \{\psi_R^*(\nu_b|X_{j,t}) - \widehat{\psi}_R^{(-\ell)}(\nu_b|X_{j,t})\}^2 = o_p(\log^{-1/2}(nT)). \tag{C.19}$$

In the following, we focus on proving (C.18). Proof of (C.19) is similar and is thus omitted.

Under (C2) and (C3), it follows from Theorem 3.7 of (Bradley, 2005) that $\{X_t\}_{t\geq 0}$ is exponentially β-mixing, that is, the β-mixing coefficient of $\{X_t\}_{t\geq 0}$ $\beta_0(\cdot)$ satisfies $\beta_0(t) = O(\rho^t)$ for some $\rho < 1$ and any $t \geq 0$. Let $n_0 = |\mathcal{I}^{(\ell)}| = n/\mathbb{L}$ and suppose $\mathcal{I}^{(\ell)} = \{\ell_1, \ell_2, \cdots, \ell_{n_0}\}$. Since $\{X_{\ell_1,t}\}_{t\geq 0}, \{X_{\ell_2,t}\}_{t\geq 0}, \cdots, \{X_{\ell_{n_0},t}\}_{t\geq 0}$ are i.i.d copies of $\{X_t\}_{t\geq 0}$, the β-mixing coefficient of

$$\{X_{\ell_1,1}, X_{\ell_1,2}, \cdots, X_{\ell_1,T}, X_{\ell_2,1}, X_{\ell_2,2}, \cdots, X_{\ell_2,T}, \cdots, X_{\ell_{n_0},1}, X_{\ell_{n_0},2}, \cdots, X_{\ell_{n_0},T}\}$$

satisfies $\beta(t) = O(\rho^t)$ for any $t \geq 0$ as well.

Let $\phi_{j,t,b}$ denote $\varphi_R^*(\mu_b|X_{j,t}) - \widehat{\varphi}_R^{(-\ell)}(\mu_b|X_{j,t})$. By (C2), we have

$$\max_{j,t,b} \mathbb{E}^{X_{j,t}} \phi_{j,t,b}^4 \leq 4 \max_{b\in\{1,\cdots,B\}} \int_x \{\varphi_R^*(\mu_b|x) - \widehat{\varphi}_R^{(-\ell)}(\mu_b|x)\}^2 \mathbb{F}(dx) \equiv \Delta, \tag{C.20}$$

where the expectation $\mathbb{E}^{X_{j,t}}$ is taken with respect to $X_{j,t}$. Notice that Δ is a random variable that depends on $\{\mu_b, \nu_b\}_{1\leq b\leq B}$ and $\{X_{j,t}\}_{j\in\mathcal{I}^{(-\ell)}, 0\leq t\leq T}$. By (C.20), we have

$$\max_{j,t,b} \mathbb{E}^{X_{j,t}} (\phi_{j,t,b}^2 - \mathbb{E}^{X_{j,t}} \phi_{j,t,b}^2)^2 \leq \Delta.$$

Under the boundedness assumption, we have $|\phi_{j,t,b}| \leq 2$ and hence $|\phi_{j,t,b}^2 - \mathbb{E}^{X_{j,t}}\phi_{j,t,b}^2| \leq 4$.

By Theorem 4.2 of Chen & Christensen (2015), we have for any integers $\tau \geq 0$ and $1 < d < n_0 T/2$ that

$$\mathbb{P}\left(\left|\sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^T (\phi_{j,t,b}^2 - \mathbb{E}^{X_0}\phi_{j,t,b}^2)\right| \geq 6\tau \,\middle|\, \Delta\right) \leq \frac{n_0 T}{d}\beta(d) + \mathbb{P}\left(\left|\sum_{(j,t)\in\mathcal{I}_r} (\phi_{j,t,b}^2 - \mathbb{E}^{X_0}\phi_{j,t,b}^2)\right|_2 \geq \tau \,\middle|\, \Delta\right)$$
$$+4\exp\left(-\frac{\tau^2/2}{n_0 T d\Delta + 4d\tau/3}\right),$$

where $\mathcal{I}_r$ denotes the last $n_0 T - d\lfloor n_0 T/d \rfloor$ elements in the list

$$\{(\ell_1, 1), (\ell_1, 2), \cdots, (\ell_1, T), (\ell_2, 1), (\ell_2, 2), \cdots, (\ell_2, T), \cdots, (\ell_{n_0}, 1), (\ell_{n_0}, 2), \cdots, (\ell_{n_0}, T)\}, \tag{C.21}$$

and $\lfloor z \rfloor$ denote the largest integer that is smaller than or equal to z for any z. Suppose $\tau \geq 4d$. Notice that $|\mathcal{I}_r| \leq d$. It follows that

$$\mathbb{P}\left(\left|\sum_{(j,t)\in\mathcal{I}_r} (\phi_{j,t,b}^2 - \mathbb{E}^{X_0}\phi_{j,t,b}^2)\right|_2 \geq \tau \,\middle|\, \Delta\right) = 0.$$

Notice that $\beta(t) = O(\rho^t)$. Set $d = -(c^* + 3)\log(n_0 T)/\log\rho$, we obtain $n_0 T\beta(d)/d = O(n_0^{-2} T^{-2} B^{-1}) = O(B^{-1} Q^{-1} n^{-2} T^{-2})$, since $Q \leq T$, $B = O((nT)^{c_*})$ and $n_0 = n/\mathbb{L}$. Here, the big-O notation is uniform in $b \in \{1, \cdots, B\}$ and $q \in \{0, \cdots, Q\}$. Set $\tau = \max\{3\sqrt{\Delta n_0 T d \log(B n_0 T)}, 11 d \log(B n_0 T)\}$, we obtain that

$$\frac{\tau^2}{4} \geq 2 n_0 T d\Delta \log(B T n_0) \quad \text{and} \quad \frac{\tau^2}{4} \geq 8 d\tau \log(B T n_0)/3 \quad \text{and} \quad \tau \geq 4d,$$

as either $n \to \infty$ or $T \to \infty$. It follows that $\tau^2/(2 n_0 T d\Delta + 8 d\tau/3) \geq 2\log(B n_0 T)$ and hence

$$\max_{b\in\{1,\cdots,B\}} \max_{q\in\{0,\cdots,Q\}} \mathbb{P}\left(\left|\sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^T (\phi_{j,t,b}^2 - \mathbb{E}^{X_0}\phi_{j,t,b}^2)\right| \geq 6\tau \,\middle|\, \Delta\right) = O(B^{-1} Q^{-1} n^{-1} T^{-1}).$$

By Bonferroni's inequality, we obtain

$$\mathbb{P}\left(\max_{b\in\{1,\cdots,B\}} \max_{q\in\{0,\cdots,Q\}} \left|\sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^T (\phi_{j,t,b}^2 - \mathbb{E}^{X_0}\phi_{j,t,b}^2)\right| \geq 6\tau \,\middle|\, \Delta\right) = O(n^{-1} T^{-1}).$$

Thus, with probability $1 - O(n^{-1}T^{-1})$, we have

$$\max_{b\in\{1,\cdots,B\}} \max_{q\in\{0,\cdots,Q\}} \left| \sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^{T} (\phi_{j,t,b}^2 - \mathbb{E}^{X_0}\phi_{j,t,b}^2) \right| = O(\sqrt{\Delta n_0 T}\log(Bn_0T), \log^2(Bn_0T)). \tag{C.22}$$

Under the given conditions on Q, we have $T-q-1$ is proportional to T for any $q \le Q$. Combining (C4) and the condition on B with (C.22) yields (C.18).

Proof of (C.15) **with** $m = 2$**:** Similar to the proof of (C.16), it suffices to show $\max_{q,b}\sqrt{n(T-q-1)}|R_{2,\ell}(q,\mu_b,\nu_b)| = o_p(\log^{-1/2}(nT))$, or $\max_{q,b}\sqrt{n(T-q-1)}|R_{2,\ell}^{(r)}(q,\mu_b,\nu_b)| = o_p(\log^{-1/2}(nT))$ for any $\ell = 1,\cdots,\mathbb{L}$ and $r = 1,2,3,4$ where

$$R_{2,\ell}(q,\mu,\nu) = \frac{1}{n(T-q-1)} \sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^{T-q-1} \{\exp(i\mu^\top S_{j,t+q+1}) - \varphi^*(\mu|X_{j,t+q})\}\{\psi^*(\nu|X_{j,t}) - \widehat{\psi}^{(-\ell)}(\nu|X_{j,t})\},$$

$$R_{2,\ell}^{(1)}(q,\mu,\nu) = \frac{1}{n(T-q-1)} \sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^{T-q-1} \{\cos(\mu^\top S_{j,t+q+1}) - \varphi_R^*(\mu|X_{j,t+q})\}\{\psi_R^*(\nu|X_{j,t}) - \widehat{\psi}_R^{(-\ell)}(\nu|X_{j,t})\},$$

$$R_{2,\ell}^{(2)}(q,\mu,\nu) = \frac{1}{n(T-q-1)} \sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^{T-q-1} \{\sin(\mu^\top S_{j,t+q+1}) - \varphi_I^*(\mu|X_{j,t+q})\}\{\psi_I^*(\nu|X_{j,t}) - \widehat{\psi}_I^{(-\ell)}(\nu|X_{j,t})\},$$

$$R_{2,\ell}^{(3)}(q,\mu,\nu) = \frac{1}{n(T-q-1)} \sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^{T-q-1} \{\cos(\mu^\top S_{j,t+q+1}) - \varphi_R^*(\mu|X_{j,t+q})\}\{\psi_I^*(\nu|X_{j,t}) - \widehat{\psi}_I^{(-\ell)}(\nu|X_{j,t})\},$$

$$R_{2,\ell}^{(4)}(q,\mu,\nu) = \frac{1}{n(T-q-1)} \sum_{j\in\mathcal{I}^{(\ell)}} \sum_{t=1}^{T-q-1} \{\sin(\mu^\top S_{j,t+q+1}) - \varphi_I^*(\mu|X_{j,t+q})\}\{\psi_R^*(\nu|X_{j,t}) - \widehat{\psi}_R^{(-\ell)}(\nu|X_{j,t})\}.$$

In the following, we only show $\max_{q,b}\sqrt{n(T-q-1)}|R_{2,\ell}^{(1)}(q,\mu_b,\nu_b)| = o_p(\log^{-1/2}(nT))$ to save space.

Define the list

$$\{(\ell_1,1),(\ell_1,2),\cdots,(\ell_1,T-q),(\ell_2,1),(\ell_2,2),\cdots,(\ell_2,T-q)\cdots,(\ell_{n_0},1),(\ell_{n_0},2),\cdots,(\ell_{n_0},T-q)\}.$$

For any $1 \le g \le n_0(T-q)$, denote by (n_g, T_g) the g-th element in the list. Let $\mathcal{F}_q^{(0)} = \{X_{\ell_1,1}, X_{\ell_1,2},\cdots,X_{\ell_1,1+q}\} \cup \{X_{j,t} : 0 \le t \le T, j \in \mathcal{I}^{(-\ell)}\} \cup \{\mu_1,\cdots,\mu_B,\nu_1,\cdots,\nu_B\}$. Then we recursively define $\mathcal{F}_q^{(g)}$ as

$$\mathcal{F}_q^{(g)} = \begin{cases} \mathcal{F}_q^{(g-1)} \cup \{X_{n_g,t_g+q+1}\}, & \text{if } g = 1 \text{ or } n_g = n_{g-1}; \\ \mathcal{F}_q^{(g-1)} \cup \{X_{n_{g-1},T}, X_{n_g,1}, X_{n_g,2},\cdots,X_{n_g,1+q}\}, & \text{otherwise.} \end{cases}$$

Let $\phi_{g,q,b}^* = \{\cos(\mu_b^\top S_{n_g,t_g+q+1}) - \varphi_R^*(\mu_b|X_{n_g,t_g+q})\}\{\psi_R^*(\nu_b|X_{n_g,t_g}) - \widehat{\psi}_R^{(-\ell)}(\nu_b|X_{n_g,t_g})\}$. Under MA, $R_{2,\ell}^{(1)}(q,\mu_b,\nu_b)$ can be rewritten as $\{n(T-q-1)\}^{-1}\sum_{g=1}^{n_0(T-q)} \phi_{g,q,b}^*$ and forms a sum of martingale difference sequence with respect to the filtration $\{\sigma(\mathcal{F}_q^{(g)}) : g \ge 0\}$ where $\sigma(\mathcal{F}_q^{(g)})$ denotes the σ-algebra generated by variables in $\mathcal{F}_q^{(g)}$. In the following, we apply concentration inequalities for martingales to bound $\max_{q,b}|R_{2,\ell}^{(1)}(q,\mu_b,\nu_b)|$.

Under the boundedness condition, we have $|\phi_{g,q,b}^*|^2 \le 4\{\psi_R^*(\nu_b|X_{n_g,t_g}) - \widehat{\psi}_R^{(-\ell)}(\nu_b|X_{n_g,t_g})\}^2$. In addition, we have by MA that

$$\begin{aligned} &\mathbb{E}\{(\phi_{g+1,q,b}^*)^2|\sigma(\mathcal{F}_q^{(g)})\} = \mathbb{E}[\{\cos(\mu_b^\top S_{n_g,t_g+q+1}) - \varphi_R^*(\mu_b|X_{n_g,t_g+q})\}^2|X_{n_g,t_g+q}] \\ &\times\{\psi_R^*(\nu_b|X_{n_g,t_g}) - \widehat{\psi}_R^{(-\ell)}(\nu_b|X_{n_g,t_g})\}^2 \le 4\{\psi_R^*(\nu_b|X_{n_g,t_g}) - \widehat{\psi}_R^{(-\ell)}(\nu_b|X_{n_g,t_g})\}^2. \end{aligned}$$

It follows from Theorem 2.1 of Bercu & Touati (2008) that

$$\mathbb{P}\left(\left|\sum_{g=1}^{n_0(T-q)} \phi_{g,q,b}^*\right| \ge \tau, \sum_{g=1}^{n_0(T-q)} 4\{\psi_R^*(\nu_b|X_{n_g,t_g}) - \widehat{\psi}_R^{(-\ell)}(\nu_b|X_{n_g,t_g})\}^2 \le y\right) \le 2\exp\left(-\frac{\tau^2}{2y}\right), \quad \forall y,\tau,$$

and hence

$$\mathbb{P}\left(\left|\sum_{g=1}^{n_0(T-q)}\phi^*_{g,q,b}\right|\geq\tau,\max_{b\in\{1,\cdots,B\}}\sum_{j\in\mathcal{I}^{(\ell)}}\sum_{t=1}^{T}\{\psi^*_R(\nu_b|X_{j,t})-\widehat{\psi}^{(-\ell)}_R(\nu_b|X_{j,t})\}^2\leq\frac{y}{4}\right)\leq 2\exp\left(-\frac{\tau^2}{2y}\right),\quad\forall y,\tau,$$

By Bonferroni's inequality, we obtain

$$\mathbb{P}\left(\max_{\substack{q\in\{0,\cdots,Q\}\\ b\in\{1,\cdots,B\}}}\left|\sum_{g=1}^{n_0(T-q)}\phi^*_{g,q,b}\right|\geq\tau,\max_{b\in\{1,\cdots,B\}}\sum_{j\in\mathcal{I}^{(\ell)}}\sum_{t=1}^{T}\{\psi^*_R(\nu_b|X_{j,t})-\widehat{\psi}^{(-\ell)}_R(\nu_b|X_{j,t})\}^2\leq\frac{y}{4}\right)\leq 2BQ\exp\left(-\frac{\tau^2}{2y}\right),$$

for any y,τ. Set $y=4\varepsilon\sqrt{nT}$, we obtain

$$\mathbb{P}\left(\max_{\substack{q\in\{0,\cdots,Q\}\\ b\in\{1,\cdots,B\}}}\left|\sum_{g=1}^{n_0(T-q)}\phi^*_{g,q,b}\right|\geq\tau,\max_{b\in\{1,\cdots,B\}}\sum_{j\in\mathcal{I}^{(\ell)}}\sum_{t=1}^{T}\{\psi^*_R(\nu_b|X_{j,t})-\widehat{\psi}^{(-\ell)}_R(\nu_b|X_{j,t})\}^2\leq\sqrt{nT}\right)$$
$$\leq 2BQ\exp\left(-\frac{\tau^2}{2\sqrt{nT}}\right),$$

It follows from (C.19) that

$$\mathbb{P}\left(\max_{\substack{q\in\{0,\cdots,Q\}\\ b\in\{1,\cdots,B\}}}\left|\sum_{g=1}^{n_0(T-q)}\phi^*_{g,q,b}\right|\geq\tau\right)\leq 2BQ\exp\left(-\frac{\tau^2}{2\sqrt{nT}}\right)+o(1). \tag{C.23}$$

Set $\tau=(nT)^{1/4}\sqrt{2\log(BQnT)}$, the right-hand-side (RHS) of (C.23) is $o(1)$. Under the given conditions on B and Q, we obtain $\max_{q,b}\sqrt{n(T-q-1)}|R^{(1)}_{2,\ell}(q,\mu_b,\nu_b)|=o_p(\log^{-1/2}(nT))$.

C.3.2. STEP 2

For any $j\in\mathcal{I}^{(\ell)}$ and $0<t<T-q$, define vectors $\lambda^*_{R,q,j,t},\lambda^*_{I,q,j,t}\in\mathbb{R}^{\mathbb{B}}$ such that the b-th element of $\lambda^*_{R,q,j,t},\lambda^*_{I,q,j,t}$ correspond to the real and imaginary part of

$$\frac{1}{\sqrt{n(T-q-1)}}\{\exp(i\mu_b^\top S_{j,t+q+1})-\varphi^*(\mu_b|X_{j,t+q})\}\{\exp(i\nu_b^\top X_{j,t-1})-\psi^*(\nu_b|X_{j,t})\},$$

respectively. Let $\lambda^*_{q,j,t}$ denote the (2B)-dimensional vector $(\lambda^{*\top}_{R,q,j,t},\lambda^{*\top}_{I,q,j,t})^\top$. In addition, we define a (2B(Q+1))-dimensional vector $\lambda^*_{j,t}$ as $(\lambda^{*\top}_{0,j,t},\lambda^{*\top}_{1,j,t-1}\mathbb{I}(t>1),\cdots,\lambda^{*\top}_{Q,j,t-Q}\mathbb{I}(t>Q))^\top$. Define the list

$$(1,1),(1,2),\cdots,(1,T-1),(2,1),(2,2),\cdots,(2,T-1),\cdots,(n,1),(n,2),\cdots,(n,T-1). \tag{C.24}$$

For any $1\leq g\leq n(T-1)$, let (n_g,t_g) be the g-th element in the list. Let $\mathcal{F}^{(0)}=\{X_{1,0}\}\cup\{\mu_1,\cdots,\mu_B,\nu_1,\cdots,\nu_B\}$ and recursively define $\mathcal{F}^{(g)}$ as

$$\mathcal{F}^{(g)}=\begin{cases}\mathcal{F}^{(g-1)}\cup\{X_{n_g,t_g}\}, & \text{if } g=1 \text{ or } n_g=n_{g-1};\\ \mathcal{F}^{(g-1)}\cup\{X_{n_{g-1},T},X_{n_g,0}\}, & \text{otherwise.}\end{cases}$$

The high-dimensional vector $M_{n,T}=\sum_{g=1}^{n(T-1)}\lambda^*_{n_g,t_g}$ forms a sum of martingale difference sequence with respect to the filtration $\{\sigma(\mathcal{F}^{(g)}):g\geq 0\}$. Notice that $S^*=\|\sum_{g=1}^{n(T-1)}\lambda^*_{n_g,t_g}\|_\infty$. In this step, we apply the high-dimensional martingale central limit theorem developed by Belloni & Oliveira (2018) to establish the limiting distribution of S^*.

For $1\leq g\leq n(T-1)$, let

$$\Sigma_g=\sum_{g=1}^{n(T-1)}\mathbb{E}\left(\lambda^*_{n_g,t_g}\lambda^{*\top}_{n_g,t_g}\middle|\mathcal{F}^{(g-1)}\right).$$

Let $V^* = \sum_{g=1}^{n(T-1)} \Sigma_g$. Using similar arguments in proving (C.22), we can show $\|V^* - V_0\|_{\infty,\infty} = O((nT)^{-1/2}\log(BnT)) + O((nT)^{-1}\log^2(BnT))$, with probability $1-O(n^{-1}T^{-1})$, where $V_0 = \mathbb{E}V^*$. Under the given conditions on B, we have $\|V^* - V_0\|_{\infty,\infty} \le \kappa_{B,n,T}$ for some $\kappa_{B,n,T} = O((nT)^{-1/2}\log(nT))$, with probability $1-O(n^{-1}T^{-1})$.

In addition, under the boundedness assumption in (C4), all the elements in V^* and V_0 are uniformly bounded by some constants. It follows that

$$\mathbb{E}\|V^* - V_0\|_{\infty,\infty} \le \kappa_{B,n,T} + \mathbb{P}(\|V^* - V_0\|_{\infty,\infty} > \kappa_{B,n,T}) = O((nT)^{-1/2}\log(nT)).$$

By Theorem 3.1 of Belloni & Oliveira (2018), we have for any Borel set $\mathcal{R}$ and any $\delta > 0$ that

$$\begin{aligned}\mathbb{P}(S^* \in \mathcal{R}) &\le \mathbb{P}(\|N(0,V_0)\|_\infty \in \mathcal{R}^{C\delta})| \\ &\le C\left(\frac{1}{nT} + \frac{\log(BnT)\log(BQ)}{\delta^2\sqrt{nT}} + \frac{\log^3(BQ)}{\delta^3\sqrt{nT}} + \frac{\log^3(BQ)}{\delta^3}\sum_{g=1}^{n(T-1)}\mathbb{E}\|\eta_g\|_\infty^3\right),\end{aligned} \tag{C.25}$$

for some constant $C > 0$.

Under the boundedness assumption in (C4), the absolute value of each element in Σ_g is uniformly bounded by $16(n(T-q-1))^{-1} = O(n^{-1}T^{-1})$. With some calculations, we can show that $\sum_{g=1}^{n(T-1)}\mathbb{E}\|\eta_g\|_\infty^3 = O((nT)^{-1/2}\log^{3/2}(BQ))$. In addition, we have $Q = O(T)$ and $B = O((nT)^{c_*})$. Combining these together with (C.25) yields

$$\mathbb{P}(S^* \in \mathcal{R}) \le \mathbb{P}(\|N(0,V_0)\|_\infty \in \mathcal{R}^{C\delta})| + O(1)\left(\frac{1}{nT} + \frac{\log^2(nT)}{\delta^2\sqrt{nT}} + \frac{\log^{9/2}(nT)}{\delta^3\sqrt{nT}}\right), \tag{C.26}$$

where $O(1)$ denotes some positive constant.

Set $\mathcal{R} = (z, +\infty)$ and $\delta = \varepsilon\log^{-1/2}(nT)/C$, we obtain

$$\mathbb{P}(S^* \le z) \ge \mathbb{P}(\|N(0,V_0)\|_\infty \le z - \varepsilon\log^{-1/2}(nT)) - o(1).$$

Set $\mathcal{R} = (-\infty, z]$, we can similarly show

$$\mathbb{P}(S^* \le z) \le \mathbb{P}(\|N(0,V_0)\|_\infty \le z + \varepsilon\log^{-1/2}(nT)) + o(1).$$

This completes the proof of Step 2.

C.3.3. Step 3

We break the proof into two parts. In Part 1, we show V_0 is a block diagonal matrix. Specifically, let V_{0,q_1,q_2} denote the $(2B)\times(2B)$ submatrix of V_0 formed by rows in $\{2q_1B+1, 2q_1B+2, \cdots, 2(q_1+1)B\}$ and columns in $\{2q_2B+1, 2q_2B+2, \cdots, 2(q_2+1)B\}$. For any $q_1 \ne q_2$, we show $V_{0,q_1,q_2} = O_{(2B)\times(2B)}$.

Let $\Sigma^{(q)}$ denote $V_{0,q,q}$. In Part 2, we provide an upper bound for $\max_{q\in\{0,\cdots,Q\}}\|\Sigma^{(q)} - \widehat{\Sigma}^{(q)}\|_{\infty,\infty}$. Let $\widehat{V}$ be a block diagonal matrix where the main diagonal blocks are given by $\widehat{\Sigma}^{(0)}, \widehat{\Sigma}^{(1)}, \cdots, \widehat{\Sigma}^{(Q)}$, we obtain $\|V_0 - \widehat{V}\|_{\infty,\infty}$

Part 1: Let $\lambda^*_{R,q,j,t,b}$ and $\lambda^*_{I,q,j,t,b}$ denote the b-th element of $\lambda^*_{R,q,j,t}$ and $\lambda^*_{I,q,j,t}$, respectively. Each element in V_{0,q_1,q_2} equals $\mathbb{E}(\sum_{j,t}\lambda^*_{Z_1,q_1,j,t,b_1})(\sum_{j,t}\lambda^*_{Z_2,q_2,j,t,b_2})$ for some $b_1, b_2 \in \{1,\cdots,B\}$ and $Z_1, Z_2 \in \{R, I\}$. In the following, we show

$$\mathbb{E}\left(\sum_{j,t}\lambda^*_{R,q_1,j,t,b_1}\right)\left(\sum_{j,t}\lambda^*_{R,q_2,j,t,b_2}\right) = 0, \quad \forall q_1 \ne q_2.$$

Similarly, one can show $\mathbb{E}(\sum_{j,t}\lambda^*_{R,q_1,j,t,b_1})(\sum_{j,t}\lambda^*_{I,q_2,j,t,b_2}) = 0$ and $\mathbb{E}(\sum_{j,t}\lambda^*_{I,q_1,j,t,b_1})(\sum_{j,t}\lambda^*_{I,q_2,j,t,b_2}) = 0$ for any $q_1 \ne q_2$. This completes the proof for Part 1.

Since observations in different trajectories are i.i.d, it suffices to show

$$\sum_j \mathbb{E}\left(\sum_t \lambda^*_{R,q_1,j,t,b_1}\right)\left(\sum_t \lambda^*_{R,q_2,j,t,b_2}\right) = 0, \qquad \forall q_1 \neq q_2,$$

or equivalently,

$$\mathbb{E}\left(\sum_t \lambda^*_{R,q_1,0,t,b_1}\right)\left(\sum_t \lambda^*_{R,q_2,0,t,b_2}\right) = 0, \qquad \forall q_1 \neq q_2, \tag{C.27}$$

By definition, we have

$$\lambda^*_{R,q,0,t,b} = \frac{1}{\sqrt{n(T-q-1)}}\{\cos(\mu_b^\top S_{t+q+1}) - \varphi^*_R(\mu_b|X_{t+q})\}\{\cos(\nu_b^\top X_{t-1}) - \psi^*_R(\nu_b|X_t)\}.$$

Since $q_1 \neq q_2$, for any t_1, t_2, we have either $t_1+q_1 \neq t_2+q_2$ or $t_1 \neq t_2$. Suppose $t_1+q_1 > t_2+q_2$. Under MA, we have

$$\mathbb{E}[\{\cos(\mu_b^\top S_{t_1+q_1+1}) - \varphi^*_R(\mu_b|X_{t_1+q_1})\}|\{X_j\}_{j\le t_1+q_1}] = 0, \qquad \forall b,$$

and hence

$$\mathbb{E}\lambda^*_{R,q_1,0,t_1,b_1}\lambda^*_{R,q_2,0,t_2,b_2} = 0, \qquad \forall b_1, b_2. \tag{C.28}$$

Similarly, when $t_1+q_1 < t_2+q_2$, we can show (C.28) holds as well.

Suppose $t_1 < t_2$, under (C1) and H_0, we have

$$\mathbb{E}[\{\cos(\nu_b^\top X_{t_1-1}) - \varphi^*_R(\nu_b|X_{t_1})\}|\{X_j\}_{j\ge t_1}] = 0, \qquad \forall b,$$

and hence (C.28) holds. Similarly, when $t_1 > t_2$, we can show (C.28) holds as well. This yields (C.27).

Part 2: For any $q \in \{0,\cdots,Q\}$, we can represent $\widehat{\Sigma}^{(q)} - \Sigma^{(q)}$ by

$$\sum_{\ell=1}^{\mathbb{L}}\sum_{j\in\mathcal{I}^{(\ell)}}\sum_{t=1}^{T-q-1}\frac{(\lambda^\top_{R,q,j,t},\lambda^\top_{I,q,j,t})^\top(\lambda^\top_{R,q,j,t},\lambda^\top_{I,q,j,t}) - (\lambda^{*\top}_{R,q,j,t},\lambda^{*\top}_{I,q,j,t})^\top(\lambda^{*\top}_{R,q,j,t},\lambda^{*\top}_{I,q,j,t})}{n(T-q-1)}. \tag{C.29}$$

Using similar arguments in Step 1 of the proof, we can show with probability tending to 1 that the absolute value of each element in (C.29) is upper bounded by $c_0^*(nT)^{-c^{**}}$ for any $q \in \{0,\cdots,Q\}$ and some positive constants $c_0, c^* > 0$. Thus we obtain $\max_{q\in\{0,\cdots,Q\}}\|\widehat{\Sigma}^{(q)} - \Sigma^{(q)}\|_{\infty,\infty} = O((nT)^{-c^{**}})$, with probability tending to 1. The proof is hence completed.

References

Belloni, A. and Oliveira, R. I. A high dimensional central limit theorem for martingales, with applications to context tree models. *arXiv preprint arXiv:1809.02741*, 2018.

Bercu, B. and Touati, A. Exponential inequalities for self-normalized martingales with applications. *Ann. Appl. Probab.*, 18 (5):1848–1869, 2008. ISSN 1050-5164. doi: 10.1214/07-AAP506.

Bradley, R. C. Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.*, 2:107–144, 2005. ISSN 1549-5787. doi: 10.1214/154957805100000104. Update of, and a supplement to, the 1986 original.

Chen, X. and Christensen, T. M. Optimal uniform convergence rates and asymptotic normality for series estimators under weak dependence and weak conditions. *J. Econometrics*, 188(2):447–465, 2015. ISSN 0304-4076. doi: 10.1016/j.jeconom.2015.03.010.

Chernozhukov, V., Chetverikov, D., and Kato, K. Detailed proof of nazarov's inequality. *arXiv preprint arXiv:1711.10696*, 2017.

Puterman, M. L. *Markov decision processes: discrete stochastic dynamic programming*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Inc., New York, 1994. ISBN 0-471-61977-9. A Wiley-Interscience Publication.