SM A

The following inequalities will be used different times in the proofs without explicit mention:

1) For any real values $a_1, \dots, a_n \geq 0$ and $s \geq 1$,

$$(a_1 + \dots + a_n)^s \leq n^{s-1}(a_1^s + \dots + a_n^s).$$

It follows immediately considering the convex function $x \mapsto x^s$ applied to the the weighted sum ${}^{a_1+\dots+a_n}/_{n}$.

2) For every values $a_1, \dots, a_n \in \mathbb{R}$ and $0 < s < 1$,

$$|a_1 + \dots + a_n|^s \leq |a_1|^s + \dots + |a_n|^s.$$

It follows immediately studying the s-Hölder function $x \mapsto |x|^s$.

By means of (2), (3) and (5), we can write for $i \geq 1$ and $l \geq 2$

$$\begin{aligned}
\varphi_{f_i^{(1)}(\mathbf{X},n)}(\mathbf{t}) &= \mathbb{E}[e^{\mathrm{i}\mathbf{t}^T f_i^{(1)}(\mathbf{X},n)}] \\
&= \mathbb{E}\Big[\exp\Big\{\mathrm{i}\mathbf{t}^T\Big[\sum_{j=1}^{I}\omega_{i,j}^{(1)}\mathbf{x}_j + b_i^{(1)}\mathbf{1}\Big]\Big\}\Big] \\
&= \mathbb{E}\Big[\exp\Big\{\mathrm{i}\mathbf{t}^T b_i^{(1)}\mathbf{1} + \mathrm{i}\mathbf{t}^T\sum_{j=1}^{I}\omega_{i,j}^{(1)}\mathbf{x}_j\Big\}\Big] \\
&= \mathbb{E}\Big[\exp\Big\{\mathrm{i}(\mathbf{t}^T\mathbf{1})b_i^{(1)}\Big\}\Big]\prod_{j=1}^{I}\mathbb{E}\Big[\exp\Big\{\mathrm{i}(\mathbf{t}^T\mathbf{x}_j)\omega_{i,j}^{(1)}\Big\}\Big] \\
&= \exp\Big\{-\frac{1}{2}\sigma_b^2(\mathbf{t}^T\mathbf{1})^2\Big\}\prod_{j=1}^{I}\exp\Big\{-\frac{1}{2}\sigma_\omega^2(\mathbf{t}^T\mathbf{x}_j)^2\Big\} \\
&= \exp\Big\{-\frac{1}{2}\Big[\sigma_b^2(\mathbf{t}^T\mathbf{1})^2 + \sigma_\omega^2\sum_{j=1}^{I}(\mathbf{t}^T\mathbf{x}_j)^2\Big]\Big\} \\
&= \exp\Big\{-\frac{1}{2}\mathbf{t}^T\Sigma(1)\mathbf{t}\Big\},
\end{aligned}$$

i.e.

$$f_i^{(1)}(\mathbf{X}) \stackrel{d}{=} N_k(\mathbf{0}, \Sigma(1)),$$

with $k \times k$ covariance matrix with element in the i-th row and j-th column as follows

$$\Sigma(1)_{i,j} = \sigma_b^2 + \sigma_\omega^2\langle x^{(i)}, x^{(j)}\rangle_{\mathbb{R}^I}.$$

Observe that we can also determine the marginal distributions,

$$f_{r,i}^{(1)}(\mathbf{X}) \sim N(0, \Sigma(1)_{r,r}), \tag{14}$$

where

$$\Sigma(1)_{r,r} = \sigma_b^2 + \sigma_\omega^2\|x^{(r)}\|_{\mathbb{R}^I}^2.$$

Now, for $i \geq 1$ and $l \geq 2$, by means of (2), (3) and (6) we can write

$$\begin{aligned}
\varphi_{f_i^{(l)}(\mathbf{X},n)|f_{1,\dots,n}^{(l-1)}}(\mathbf{t}) &= \mathbb{E}[e^{\mathrm{i}\mathbf{t}^T f_i^{(l)}(\mathbf{X},n)}|f_{1,\dots,n}^{(l-1)}] \\
&= \mathbb{E}\Big[\exp\Big\{\mathrm{i}\mathbf{t}^T\Big[\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\omega_{i,j}^{(l)}(\phi \bullet f_j^{(l-1)}(\mathbf{X},n)) + b_i^{(1)}\mathbf{1}\Big]\Big\}|f_{1,\dots,n}^{(l-1)}\Big] \\
&= \mathbb{E}\Big[\exp\Big\{\mathrm{i}\mathbf{t}^T b_i^{(1)}\mathbf{1} + \mathrm{i}\mathbf{t}^T\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\omega_{i,j}^{(l)}(\phi \bullet f_j^{(l-1)}(\mathbf{X},n))\Big\}|f_{1,\dots,n}^{(l-1)}\Big] \\
&= \mathbb{E}\Big[\exp\Big\{\mathrm{i}(\mathbf{t}^T\mathbf{1})b_i^{(1)}\Big\}\Big]\prod_{j=1}^{n}\mathbb{E}\Big[\exp\Big\{\mathrm{i}\omega_{i,j}^{(l)}\Big(\frac{1}{\sqrt{n}}\mathbf{t}^T(\phi \bullet f_j^{(l-1)}(\mathbf{X},n))\Big)\Big\}|f_{1,\dots,n}^{(l-1)}\Big] \\
&= \exp\Big\{-\frac{1}{2}\sigma_b^2(\mathbf{t}^T\mathbf{1})^2\Big\}\prod_{j=1}^{n}\exp\Big\{-\frac{1}{2n}\sigma_\omega^2\Big(\mathbf{t}^T(\phi \bullet f_j^{(l-1)}(\mathbf{X},n))\Big)^2\Big\} \\
&= \exp\Big\{-\frac{1}{2}\Big[\sigma_b^2(\mathbf{t}^T\mathbf{1})^2 + \frac{\sigma_\omega^2}{n}\sum_{j=1}^{n}\Big(\mathbf{t}^T(\phi \bullet f_j^{(l-1)}(\mathbf{X},n))\Big)^2\Big]\Big\} \\
&= \exp\Big\{-\frac{1}{2}\mathbf{t}^T\Sigma(l,n)\mathbf{t}\Big\},
\end{aligned}$$

i.e.

$$f_i^{(l)}(\mathbf{X},n)|f_{1,\dots,n}^{(l-1)} \stackrel{d}{=} N_k(\mathbf{0},\Sigma(l,n)),$$

with $k \times k$ covariance matrix with element in the i-th row and j-th column as follows

$$\Sigma(l,n)_{i,j} = \sigma_b^2 + \frac{\sigma_\omega^2}{n}\Big\langle(\phi \bullet \mathbf{F}_i^{(l-1)}(\mathbf{X},n)),(\phi \bullet \mathbf{F}_j^{(l-1)}(\mathbf{X},n))\Big\rangle_{\mathbb{R}^n}.$$

Observe that we can also determine the marginal distributions,

$$f_{r,i}^{(l)}(\mathbf{X},n)|f_{1,\dots,n}^{(l-1)} \sim N(0,\Sigma(l,n)_{r,r}), \tag{15}$$

where

$$\Sigma(l,n)_{r,r} = \sigma_b^2 + \frac{\sigma_\omega^2}{n}\|\phi \bullet \mathbf{F}_r^{(l-1)}(\mathbf{X},n)\|_{\mathbb{R}^n}^2.$$

SM A.1: ASYMPTOTICS FOR THE $i-th$ COORDINATE

First of all, from Definition 1, note that since $f_i^{(1)}(\mathbf{X})$ does not depend on n we consider the limit as $n \to \infty$ only for $f_i^{(l)}(\mathbf{X},n)$ for all $l \geq 2$. It comes directly from Equation (6) that, for every fixed l and n the sequence $\big(f_i^{(l)}(\mathbf{X},n)\big)_{i\geq 1}$ is exchangeable. In particular, let $p_n^{(l)}$ denote the de Finetti (random) probability measure of the exchangeable sequence $\big(f_i^{(l)}(\mathbf{X},n)\big)_{i\geq 1}$. That is, by the celebrated de Finetti representation theorem, conditionally to $p_n^{(l)}$ the $f_i^{(l)}(\mathbf{X},n)$'s are iid as $p_n^{(l)}$. Now, let consider the induction hypothesis that, $p_n^{(l-1)} \stackrel{d}{\to} q^{(l-1)}$ as $n \to +\infty$, where $q^{(l-1)} = N_k(\mathbf{0},\Sigma(l-1))$. To establish the convergence in distribution we rely on Theorem 5.3 of Kallenberg (2002) known as Levy theorem, taking into account the point-wise convergence of the characteristic functions. Therefore we can write the following expression:

$$
\begin{aligned}
\varphi_{f_i^{(l)}(\mathbf{X},n)}(\mathbf{t}) &= \mathbb{E}[e^{\mathrm{i}\mathbf{t}^T f_i^{(l)}(\mathbf{X},n)}] \\
&= \mathbb{E}[\mathbb{E}[e^{\mathrm{i}\mathbf{t}^T f_i^{(l)}(\mathbf{X},n)}|f_{1,\dots,n}^{(l-1)}]] \\
&= \mathbb{E}\Big[\exp\Big\{-\frac{1}{2}\mathbf{t}^T\Sigma(l,n)\mathbf{t}\Big\}\Big] \\
&= \mathbb{E}\Big[\exp\Big\{-\frac{1}{2}\Big[\sigma_b^2(\mathbf{t}^T\mathbf{1})^2 + \frac{\sigma_\omega^2}{n}\sum_{j=1}^n\Big(\mathbf{t}^T(\phi\bullet f_j^{(l-1)}(\mathbf{X},n))\Big)^2\Big]\Big\}\Big] \\
&= e^{-\frac{1}{2}\sigma_b^2(\mathbf{t}^T\mathbf{1})^2}\mathbb{E}\Big[\exp\Big\{-\frac{\sigma_\omega^2}{2n}\sum_{j=1}^n\Big(\mathbf{t}^T(\phi\bullet f_j^{(l-1)}(\mathbf{X},n))\Big)^2\Big\}\Big] \\
&= e^{-\frac{1}{2}\sigma_b^2(\mathbf{t}^T\mathbf{1})^2}\mathbb{E}\Big[\mathbb{E}\Big[\exp\Big\{-\frac{\sigma_\omega^2}{2n}\sum_{j=1}^n\Big(\mathbf{t}^T(\phi\bullet f_j^{(l-1)}(\mathbf{X},n))\Big)^2\Big\}|p_n^{(l-1)}\Big]\Big] \\
&= e^{-\frac{1}{2}\sigma_b^2(\mathbf{t}^T\mathbf{1})^2}\mathbb{E}\Big[\prod_{j=1}^n\mathbb{E}\Big[\exp\Big\{-\frac{\sigma_\omega^2}{2n}\Big(\mathbf{t}^T(\phi\bullet f_j^{(l-1)}(\mathbf{X},n))\Big)^2\Big\}|p_n^{(l-1)}\Big]\Big] \\
&= e^{-\frac{1}{2}\sigma_b^2(\mathbf{t}^T\mathbf{1})^2}\mathbb{E}\Big[\prod_{j=1}^n\int\exp\Big\{-\frac{\sigma_\omega^2}{2n}\Big(\mathbf{t}^T(\phi\bullet f)\Big)^2\Big\}p_n^{(l-1)}(\mathrm{d}f)\Big] \\
&= e^{-\frac{1}{2}\sigma_b^2(\mathbf{t}^T\mathbf{1})^2}\mathbb{E}\Big[\Big(\int\exp\Big\{-\frac{\sigma_\omega^2}{2n}\Big(\mathbf{t}^T(\phi\bullet f)\Big)^2\Big\}p_n^{(l-1)}(\mathrm{d}f)\Big)^n\Big].
\end{aligned}
$$

Observe that the last integral is with respect to k coordinates: i.e. $\mathrm{d}f = (\mathrm{d}f_1,\dots,\mathrm{d}f_k)$. Denote as $\xrightarrow{p}$ the convergence in probability. We will prove the following lemmas:

- L1) for each $l \geq 2$ and $s \geq 1$, $\mathbb{P}[p_n^{(l-1)} \in Y_s] = 1$, where $Y_s = \{p : \int \|\phi\bullet f\|_{\mathbb{R}^k}^s p(\mathrm{d}f) < +\infty\}$;
- L2) $\int(\mathbf{t}^T(\phi\bullet f))^2 p_n^{(l-1)}(\mathrm{d}f) \xrightarrow{p} \int(\mathbf{t}^T(\phi\bullet f))^2 q^{(l-1)}(\mathrm{d}f)$, as $n\to+\infty$;
- L3) $\int(\mathbf{t}^T(\phi\bullet f))^2\big[1-\exp\big\{-\theta\frac{\sigma_\omega^2}{2n}(\mathbf{t}^T(\phi\bullet f))^2\big\}\big]p_n^{(l-1)}(\mathrm{d}f) \xrightarrow{p} 0$, as $n\to+\infty$ for every $\theta\in(0,1)$.

SM A.1.1: PROOF OF L1

In order to prove the three lemmas, we will use many times the envelope condition (4) without explicit mention. For $l = 2$ we have

$$
\begin{aligned}
\mathbb{E}[\|\phi\bullet f_i^{(1)}(\mathbf{X})\|_{\mathbb{R}^k}^s] &\leq \mathbb{E}\Big[\Big(\sum_{r=1}^k|\phi\circ f_{r,i}^{(1)}(\mathbf{X})|^2\Big)^{s/2}\Big] \\
&\leq \mathbb{E}\Big[\Big(\sum_{r=1}^k|\phi\circ f_{r,i}^{(1)}(\mathbf{X})|\Big)^s\Big] \\
&\leq \mathbb{E}\Big[k^{s-1}\sum_{r=1}^k|\phi\circ f_{r,i}^{(1)}(\mathbf{X})|^s\Big] \\
&= k^{s-1}\sum_{r=1}^k\mathbb{E}\Big[|\phi\circ f_{r,i}^{(1)}(\mathbf{X})|^s\Big] \\
&\leq k^{s-1}\sum_{r=1}^k\mathbb{E}\Big[(a+b|f_{r,i}^{(1)}(\mathbf{X})|^m)^s\Big] \\
&\leq (2k)^{s-1}\sum_{r=1}^k\Big(a^s+b^s\mathbb{E}[|f_{r,i}^{(1)}(\mathbf{X})|^{sm}]\Big) \\
&< +\infty,
\end{aligned}
$$

where we used that from (14), $f_{r,i}^{(1)}(\mathbf{X}) \sim N(0, \sigma_b^2 + \sigma_\omega^2 \|x^{(r)}\|^2_{\mathbb{R}^I})$ and then

$$\mathbb{E}[|f_{r,i}^{(1)}(\mathbf{X})|^{sm}] = M_{sm}(\sigma_b^2 + \sigma_\omega^2 \|x^{(r)}\|^2_{\mathbb{R}^I})^{sm/2},$$

where M_c is the c-th moment of $|N(0,1)|$. Now assume that L1 is true for $(l-2)$, i.e. for each $s \geq 1$ it holds $\int \|\phi \bullet f\|^s_{\mathbb{R}^k} p_n^{(l-2)}(\mathrm{d}f) < +\infty$ uniformly in n, and we prove that it is true also for $(l-1)$.

$$\begin{aligned}
\mathbb{E}[\|\phi \bullet f_i^{(l-1)}(\mathbf{X}, n)\|^s_{\mathbb{R}^k} | f_{1,\dots,n}^{(l-2)}] &\leq \mathbb{E}\Big[k^{s-1}\sum_{r=1}^k |\phi \circ f_{r,i}^{(l-1)}(\mathbf{X}, n)|^s | f_{1,\dots,n}^{(l-2)}\Big] \\
&\leq (2k)^{s-1}\sum_{r=1}^k \Big(a^s + b^s \mathbb{E}\Big[|f_{r,i}^{(l-1)}(\mathbf{X}, n)|^{ms} | f_{1,\dots,n}^{(l-2)}\Big]\Big) \\
&\leq D_1(a, k.s) + D_2(b, k, s)\sum_{r=1}^k \mathbb{E}\Big[|f_{r,i}^{(l-1)}(\mathbf{X}, n)|^{ms} | f_{1,\dots,n}^{(l-2)}\Big].
\end{aligned}$$

From (15) we get

$$\begin{aligned}
\mathbb{E}\Big[|f_{r,i}^{(l-1)}(\mathbf{X}, n)|^{ms} | f_{1,\dots,n}^{(l-2)}\Big] &= M_{ms}\Big(\sigma_b^2 + \frac{\sigma_\omega^2}{n}\|\phi \bullet \mathbf{F}_r^{(l-2)}(\mathbf{X}, n)\|^2_{\mathbb{R}^n}\Big)^{sm/2} \\
&\leq M_{ms} 2^{sm-1}\Big(\sigma_b^{2sm} + \frac{\sigma_\omega^{2sm}}{n^{sm}}\|\phi \bullet \mathbf{F}_r^{(l-2)}(\mathbf{X}, n)\|^{2sm}_{\mathbb{R}^n}\Big)^{1/2}.
\end{aligned}$$

Thus we have

$$\begin{aligned}
&\mathbb{E}[\|\phi \bullet f_i^{(l-1)}(\mathbf{X}, n)\|^s_{\mathbb{R}^k} | p_n^{(l-2)}] \\
&\leq D_1(a, k, s) + D_3(b, k, s, m)\sum_{r=1}^k \Big(\sigma_b^{2sm} + \frac{\sigma_\omega^{2sm}}{n^{sm}}\mathbb{E}\Big[\|\phi \bullet \mathbf{F}_r^{(l-2)}(\mathbf{X}, n)\|^{2sm}_{\mathbb{R}^n} | p_n^{(l-2)}\Big]\Big)^{1/2},
\end{aligned}$$

where

$$\begin{aligned}
\mathbb{E}\Big[\|\phi \bullet \mathbf{F}_r^{(l-2)}(\mathbf{X}, n)\|^{2sm}_{\mathbb{R}^n} | p_n^{(l-2)}\Big] &\leq \mathbb{E}\Big[n^{sm-1}\sum_{i=1}^n |\phi \circ f_{r,i}^{(l-2)}(\mathbf{X}, n)|^{2sm} | p_n^{(l-2)}\Big] \\
&\leq D_4(s, m) n^{sm}\int |\phi(f_r)|^{2sm} p_n^{(l-2)}(\mathrm{d}f_r) \\
&\leq D_4(s, m) n^{sm}\int \|\phi \bullet f\|^{2sm}_{\mathbb{R}^k} p_n^{(l-2)}(\mathrm{d}f),
\end{aligned}$$

where the last inequality is due to the fact that $|\phi(f_r)|^{2sm} \leq \Big(\sum_{r=1}^k |\phi(f_r)|^2\Big)^{sm}$ and then

$$\int |\phi(f_r)|^{2sm} p_n^{(l-2)}(\mathrm{d}f_r) \leq \int \Big(\sum_{r=1}^k |\phi(f_r)|^2\Big)^{sm} p_n^{(l-2)}(\mathrm{d}f_1, \dots, \mathrm{d}f_k) = \int \|\phi \bullet f\|^{2sm}_{\mathbb{R}^k} p_n^{(l-2)}(\mathrm{d}f).$$

So, we proved that

$$\begin{aligned}
&\mathbb{E}[\|\phi \bullet f_i^{(l-1)}(\mathbf{X}, n)\|^s_{\mathbb{R}^k} | p_n^{(l-2)}] \\
&\leq D_1(a, k, s) + D_3(b, k, s, m)\sum_{r=1}^k \Big(\sigma_b^{2sm} + \sigma_\omega^{2sm} D_4(s, m)\int \|\phi \bullet f\|^{2sm}_{\mathbb{R}^k} p_n^{(l-2)}(\mathrm{d}f)\Big)^{1/2},
\end{aligned} \tag{16}$$

which is finite by induction hypothesis uniformly in n. To conclude, since $p_n^{(l-1)} \overset{iid}{\sim} f_i^{(l-1)}(\mathbf{X}, n) | p_n^{(l-1)}$ we get

$$\begin{aligned}
\int \|\phi \bullet f\|^s_{\mathbb{R}^k} p_n^{(l-1)}(\mathrm{d}f) &= \mathbb{E}[\|\phi \bullet f_i^{(l-1)}(\mathbf{X}, n)\|^s_{\mathbb{R}^k} | p_n^{(l-1)}] \\
&= \mathbb{E}[\mathbb{E}[\|\phi \bullet f_i^{(l-1)}(\mathbf{X}, n)\|^s_{\mathbb{R}^k} | p_n^{(l-2)}] | p_n^{(l-1)}] \\
&\leq cost(a, k, s, m) < \infty
\end{aligned} \tag{17}$$

which is bounded uniformly in n since the inner expectation is bounded uniformly in n by (16).

Remark: Y_s is a measurable set with respect to the weak topology for each $s \geq 1$, indeed for each $R \in \mathbb{N}$ defining the map

$$T_R : U \to \mathbb{R}, \quad T_R(p) = \int_{B_R(0)} \|\phi \bullet f\|^s_{\mathbb{R}^k} p(\mathrm{d}f) = \int_{\mathbb{R}^k} \|\phi \bullet f\|^s_{\mathbb{R}^k} \mathcal{X}_{(B_R(0))}(f) p(\mathrm{d}f)$$

where $U := \{p : p$ distribution of a r.v. $X : \Omega \to \mathbb{R}^k\}$ endowed with the weak topology, since $\cap_{R\in\mathbb{N}} T_R^{-1}(0,\infty) = Y_s$ and $(0,\infty)$ is open, it is sufficient to prove that T_R is continuous. Let $(p_m) \subset U$ such that p_m converges to p with respect to the weak topology, then by Definition 3

$$|T_R(p_m) - T_R(p)| = \Big| \int \|\phi \bullet f\|^s_{\mathbb{R}^k} \mathcal{X}_{(B_R(0))}(f) p_m(\mathrm{d}f) - \int \|\phi \bullet f\|^s_{\mathbb{R}^k} \mathcal{X}_{(B_R(0))}(f) p(\mathrm{d}f) \Big| \to 0$$

because the function $f \mapsto \|\phi \bullet f\|^s_{\mathbb{R}^k} \mathcal{X}_{(B_R(0))}(f)$ is continuous (by composition of the continuous functions ϕ and $\|\|^s$) and bounded by Weierstrass theorem.

SM A.1.2: PROOF OF L2

By induction hypothesis, $p_n^{(l-1)}$ converges weakly to a $p^{(l-1)}$ with respect to the weak topology and the limit is degenerate, in the sense that it provides a.s. the distribution $q^{(l-1)}$. Then $p_n^{(l-1)}$ converges in probability to $p^{(l-1)}$. Then for every sub sequence n' there exists a further sub sequence n'' such that $p_{n''}^{(l-1)}$ converges a.s. to $p^{(l-1)}$. By induction hypothesis, $p^{(l-1)}$ is absolutely continuous with respect to the Lebesgue measure. Since ϕ is a.s. continuous and the sequence $\big((\mathbf{t}^T(\phi \bullet f))^2\big)_{n\geq 1}$ uniformly integrable with respect to $p_n^{(l-1)}$ (by Cauchy-Schwarz inequality and L1 $\int \big(\mathbf{t}^T(\phi \bullet f)\big)^{2s} p_n^{(l-1)}(\mathrm{d}f) \leq \|\mathbf{t}\|^{2s}_{\mathbb{R}^k} \int \|\phi \bullet f\|^{2s}_{\mathbb{R}^k} p_n^{(l-1)}(\mathrm{d}f) < \infty$, thus is L^s-bounded for each $s \geq 1$, and so uniformly integrable, then we can write the following

$$\int (\mathbf{t}^T(\phi \bullet f))^2 p_{n''}^{(l-1)}(\mathrm{d}f) \overset{a.s.}{\to} \int (\mathbf{t}^T(\phi \bullet f))^2 q^{(l-1)}(\mathrm{d}f).$$

Thus, as $n \to +\infty$

$$\int (\mathbf{t}^T(\phi \bullet f))^2 p_n^{(l-1)}(\mathrm{d}f) \overset{p}{\to} \int (\mathbf{t}^T(\phi \bullet f))^2 q^{(l-1)}(\mathrm{d}f).$$

SM A.1.3: PROOF OF L3

Let $p \geq 1$ and $q \geq 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$. By means of Hölder inequality

$$\begin{aligned}
&\int \|\phi \bullet f\|^2_{\mathbb{R}^k} (1 - e^{-\frac{\sigma^2_\omega}{2n}(\mathbf{t}^T(\phi\bullet f))^2}) p_n^{(l-1)}(\mathrm{d}f) \\
&\leq \Big(\int \|\phi \bullet f\|^{2p}_{\mathbb{R}^k} p_n^{(l-1)}(\mathrm{d}f) \Big)^{1/p} \Big(\int (1 - e^{-\frac{\sigma^2_\omega}{2n}(\mathbf{t}^T(\phi\bullet f))^2})^q p_n^{(l-1)}(\mathrm{d}f) \Big)^{1/q}.
\end{aligned}$$

Since $q \geq 1$, for every $y \geq 0$ we have $0 \leq 1 - e^{-y} < 1$, then $(1-e^{-y})^q \leq (1 - e^{-y}) \leq y$. It implies the following

$$\begin{aligned}
&\int \|\phi \bullet f\|^2_{\mathbb{R}^k} (1 - e^{-\frac{\sigma^2_\omega}{2n}(\mathbf{t}^T(\phi\bullet f))^2}) p_n^{(l-1)}(\mathrm{d}f) \\
&\leq \Big(\int \|\phi \bullet f\|^{2p}_{\mathbb{R}^k} p_n^{(l-1)}(\mathrm{d}f) \Big)^{1/p} \Big(\int \frac{\sigma^2_\omega}{2n} (\mathbf{t}^T(\phi \bullet f))^2 p_n^{(l-1)}(\mathrm{d}f) \Big)^{1/q} \\
&\leq \Big(\int \|\phi \bullet f\|^{2p}_{\mathbb{R}^k} p_n^{(l-1)}(\mathrm{d}f) \Big)^{1/p} \Big(\|\mathbf{t}\|^2_{\mathbb{R}^k} \frac{\sigma^2_\omega}{2n} \int \|\phi \bullet f\|^2_{\mathbb{R}^k} p_n^{(l-1)}(\mathrm{d}f) \Big)^{1/q} \to 0,
\end{aligned}$$

as $n \to +\infty$ since by L1 the two integrals are bounded uniformly in n. Thus for every $y > 0$ and $\theta \in (0,1)$ $e^{-\theta y} \geq e^{-y} \Rightarrow 0 \leq 1 - e^{-\theta y} \leq 1 - e^{-y} \leq 1$ we get

$$\begin{aligned}
0 &\leq \int \big(\mathbf{t}^T(\phi \bullet f)\big)^2 \Big[1 - \exp\Big\{ -\theta \frac{\sigma^2_\omega}{2n} \big(\mathbf{t}^T(\phi \bullet f)\big)^2 \Big\}\Big] p_n^{(l-1)}(\mathrm{d}f) \\
&\leq \|\mathbf{t}\|^2_{\mathbb{R}^k} \int \|\phi \bullet f\|^2_{\mathbb{R}^k} \Big[1 - \exp\Big\{ - \frac{\sigma^2_\omega}{2n} \big(\mathbf{t}^T(\phi \bullet f)\big)^2 \Big\}\Big] p_n^{(l-1)}(\mathrm{d}f) \to 0,
\end{aligned}$$

as $n \to +\infty$.

SM A.1.4: COMBINATION OF THE LEMMAS

We conclude in two steps.

Step 1: uniform integrability. Define $y = y_n(f) = \frac{\sigma_\omega^2}{2n}(\mathbf{t}^T(\phi \bullet f))^2$. Thus

$$\begin{aligned}\varphi_{f_i^{(l)}(\mathbf{X},n)}(\mathbf{t}) &= e^{-\frac{1}{2}\sigma_b^2(\mathbf{t}^T\mathbf{1})^2}\mathbb{E}\Big[\Big(\int e^{-y_n(f)}p_n^{(l-1)}(\mathrm{d}f)\Big)^n\Big] \\ &= e^{-\frac{1}{2}\sigma_b^2(\mathbf{t}^T\mathbf{1})^2}\mathbb{E}[A_n]\end{aligned}$$

where $A_n = \Big(\int e^{-y_n(f)}p_n^{(l-1)}(\mathrm{d}f)\Big)^n$. $(A_n)_{n\geq 1}$ is is uniformly integrable because it is L^s-bounded for all $s \geq 1$. Indeed, since $0 < e^{-y_n(f)} \leq 1$

$$\mathbb{E}[A_n^s] \leq \mathbb{E}\Big[\Big(\int p_n^{(l-1)}(\mathrm{d}f)\Big)^{ns}\Big] = \mathbb{E}\big[1\big] = 1$$

Step 2: convergence in probability. By Lagrange theorem for $y > 0$ there exists $\theta \in (0,1)$ such that $e^{-y} = 1 - y + y(1 - e^{-y\theta})$. Then for every n there exists a real value $\theta_n \in (0,1)$ such that the follow equality holds:

$$A_n = \Big(1 - \frac{\sigma_\omega^2}{2n}[A_n' - A_n'']\Big)^n.$$

where

$$\begin{cases} A_n' = \int(\mathbf{t}^T(\phi \bullet f))^2 p_n^{(l-1)}(\mathrm{d}f) \\ A_n'' = \int(\mathbf{t}^T(\phi \bullet f))^2\Big[1 - \exp\Big\{-\theta_n\frac{\sigma_\omega^2}{2n}(\mathbf{t}^T(\phi \bullet f))^2\Big\}\Big]p_n^{(l-1)}(\mathrm{d}f)\end{cases}$$

Using the definition of the exponential function, i.e. $e^x = \lim_{n\to\infty}(1 + \frac{x}{n})^n$, L2 and L3 we get that

$$A_n \xrightarrow{p} \exp\Big\{-\frac{\sigma_\omega^2}{2}\int(\mathbf{t}^T(\phi \bullet f))^2 q^{(l-1)}(\mathrm{d}f)\Big\}, \quad \text{as } n \to \infty$$

Conclusion: since convergence in probability with uniform integrability implies convergence in mean, by the two above steps we get

$$\begin{aligned}\varphi_{f_i^{(l)}(\mathbf{X},n)}(\mathbf{t}) = e^{-\frac{1}{2}\sigma_b^2(\mathbf{t}^T\mathbf{1})^2}\mathbb{E}[A_n] &\to \exp\Big\{-\frac{\sigma_b^2}{2}(\mathbf{t}^T\mathbf{1})^2 - \frac{\sigma_\omega^2}{2}\int(\mathbf{t}^T(\phi \bullet f))^2 q^{(l-1)}(\mathrm{d}f)\Big\} \\ &= \exp\Big\{-\frac{1}{2}\Big[\sigma_b^2(\mathbf{t}^T\mathbf{1})^2 + \sigma_\omega^2\int(\mathbf{t}^T(\phi \bullet f))^2 q^{(l-1)}(\mathrm{d}f)\Big]\Big\} \\ &= \exp\Big\{-\frac{1}{2}\mathbf{t}^T\Sigma(l)\mathbf{t}\Big\},\end{aligned}$$

where $\Sigma(l)$ is a $k \times k$ matrix with elements

$$\Sigma(l)_{i,j} = \sigma_b^2 + \sigma_\omega^2\int \phi(f_i)\phi(f_j)q^{(l-1)}(\mathrm{d}f),$$

where $q^{(l-1)} = N_k(\mathbf{0}, \Sigma(l-1))$. Then the limit distribution of $f_i^{(l)}(\mathbf{X})$ is a k-dimensional Gaussian distribution with mean $\mathbf{0}$ and covariance matrix $\Sigma(l)$, i.e. as $n \to +\infty$,

$$f_i^{(l)}(\mathbf{X}, n) \xrightarrow{d} N_k(\mathbf{0}, \Sigma(l)).$$

SM B.1

Fix $i \geq 1$, $l \geq 1$, $n \in \mathbb{N}$. We prove that there exists a random variable $H_i^{(l)}(n)$ such that

$$|f_i^{(l)}(x,n) - f_i^{(l)}(y,n)| \leq H_i^{(l)}(n)\|x - y\|_{\mathbb{R}^I}, \quad x, y \in \mathbb{R}^I, \mathbb{P} - a.s.$$

i.e. fixed $\xi \in \Omega$ the function $x \mapsto f_i^{(l)}(x,n)(\xi)$ is Lipschitz. We proceed by induction on the layers. Fix $x, y \in \mathbb{R}^I$. For the first layer, by (5) we get

$$
\begin{aligned}
|f_i^{(1)}(x,n)(\xi) - f_i^{(1)}(y,n)(\xi)| &= \Big| \sum_{j=1}^{I} \omega_{i,j}^{(1)}(\xi) x_j + b_i^{(1)}(\xi) - \big(\sum_{j=1}^{I} \omega_{i,j}^{(1)}(\xi) y_j + b_i^{(1)}(\xi) \big) \Big| \\
&= \Big| \sum_{j=1}^{I} \omega_{i,j}^{(1)}(\xi) x_j - \sum_{j=1}^{I} \omega_{i,j}^{(1)}(\xi) y_j \Big| \\
&= \Big| \sum_{j=1}^{I} \omega_{i,j}^{(1)}(\xi)(x_j - y_j) \Big| \\
&\leq \sum_{j=1}^{I} \Big| \omega_{i,j}^{(1)}(\xi) \Big| |x_j - y_j| \\
&\leq \|x - y\|_{\mathbb{R}^I} \sum_{j=1}^{I} \Big| \omega_{i,j}^{(1)}(\xi) \Big|
\end{aligned}
$$

where we used that $|x_j - y_j| \leq \|x - y\|_{\mathbb{R}^I}$. Set $H_i^{(1)}(n) = \sum_{j=1}^{I} \big| \omega_{i,j}^{(1)} \big|$. Suppose by induction hypothesis that for each $j \geq 1$ there exists a random variable $H_j^{(l-1)}(n)$ such that $|f_j^{(l-1)}(x,n)(\xi) - f_j^{(l-1)}(y,n)(\xi)| \leq H_j^{(l-1)}(n)(\xi) \|x - y\|_{\mathbb{R}^I}$, and let L_ϕ be the Lipschitz constant of ϕ. Then by (6) we get

$$
\begin{aligned}
&|f_i^{(l)}(x,n)(\xi) - f_i^{(l)}(y,n)(\xi)| \\
&\quad = \Big| \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \omega_{i,j}^{(l)}(\xi) \phi(f_j^{(l-1)}(x,n)) + b_i^{(l)}(\xi) - \Big[\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \omega_{i,j}^{(l)}(\xi) \phi(f_j^{(l-1)}(y,n)) + b_i^{(l)}(\xi) \Big] \Big| \\
&\quad = \Big| \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \omega_{i,j}^{(l)}(\xi) \phi(f_j^{(l-1)}(x,n)) - \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \omega_{i,j}^{(l)}(\xi) \phi(f_j^{(l-1)}(y,n)) \Big| \\
&\quad \leq \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \Big| \omega_{i,j}^{(l)}(\xi) \Big| \Big| \phi(f_j^{(l-1)}(x,n)) - \phi(f_j^{(l-1)}(y,n)) \Big| \\
&\quad \leq \frac{L_\phi}{\sqrt{n}} \sum_{j=1}^{n} \Big| \omega_{i,j}^{(l)}(\xi) \Big| \Big| f_j^{(l-1)}(x,n) - f_j^{(l-1)}(y,n) \Big| \\
&\quad \leq \frac{L_\phi}{\sqrt{n}} \sum_{j=1}^{n} \Big| \omega_{i,j}^{(l)}(\xi) \Big| H_j^{(l-1)}(n)(\xi) \|x - y\|_{\mathbb{R}^I} \\
&\quad \leq \|x - y\|_{\mathbb{R}^I} \frac{L_\phi}{\sqrt{n}} \sum_{j=1}^{n} \Big| \omega_{i,j}^{(l)}(\xi) \Big| H_j^{(l-1)}(n)(\xi)
\end{aligned}
$$

Set

$$
H_i^{(l)}(n) = \frac{L_\phi}{\sqrt{n}} \sum_{j=1}^{n} \Big| \omega_{i,j}^{(l)} \Big| H_j^{(l-1)}(n)
$$

Thus we proved that fixed $l \geq 1$, and $i \geq 1$, for each $n \in \mathbb{N}$

$$
\mathbb{P}\Big[\Big\{ \xi \in \Omega : |f_i^{(l)}(x,n)(\xi) - f_i^{(l)}(y,n)(\xi)| \leq H_i^{(l)}(n)(\xi) \|x - y\|_{\mathbb{R}^I} \Big\} \Big] = 1.
$$

Thus, each process $f_i^{(l)}(1), f_i^{(l)}(2), \ldots$ is $\mathbb{P}$-a.s. Lipschitz, in particular is $\mathbb{P}$-a.s. continuous processes, i.e. belongs to $C(\mathbb{R}^I; \mathbb{R})$. In order to prove the continuity of $f_i^{(l)}$ we can not just take the limit as $n \to +\infty$ of (9) because the left quantity converges to $|f_i^{(l)}(x) - f_i^{(l)}(y)|$ only in distribution and not $\mathbb{P}$-a.s., but we can prove the continuity by applying Proposition 2, as we will show in SM B.2.

SM B.2

Fix $i \geq 1, l \geq 1$. We show the continuity of the limiting process $f_i^{(l)}$ by applying Proposition 2. Take two inputs $x, y \in \mathbb{R}^I$. From (7) we know that $[f_i^{(l)}(x), f_i^{(l)}(y)] \sim N_2(\mathbf{0}, \Sigma(l))$ where

$$\Sigma(1) = \sigma_b^2 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \sigma_\omega^2 \begin{bmatrix} \|x\|^2_{\mathbb{R}^I} & \langle x, y\rangle_{\mathbb{R}^I} \\ \langle x, y\rangle_{\mathbb{R}^I} & \|y\|^2_{\mathbb{R}^I} \end{bmatrix},$$
$$\Sigma(l) = \sigma_b^2 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \sigma_\omega^2 \int \begin{bmatrix} |\phi(u)|^2 & \phi(u)\phi(v) \\ \phi(u)\phi(v) & |\phi(v)|^2 \end{bmatrix} q^{(l-1)}(\mathrm{d}u, \mathrm{d}v),$$

where $q^{(l-1)} = N_2(\mathbf{0}, \Sigma(l-1))$. We want to find two values $\alpha > 0$ and $\beta > 0$, and a constant $H^{(l)} > 0$ such that

$$\mathbb{E}\Big[|f_i^{(l)}(y) - f_i^{(l)}(x)|^\alpha\Big] \leq H^{(l)} \|y - x\|^{I+\beta}_{\mathbb{R}^I}.$$

Defining $\mathbf{a}^T = [1, -1]$ we have $f_i^{(l)}(y) - f_i^{(l)}(x) \sim N(\mathbf{a}^T\mathbf{0}, \mathbf{a}^T\Sigma(l)\mathbf{a})$. Consider $\alpha = 2\theta$ with θ integer. Thus

$$|f_i^{(l)}(y) - f_i^{(l)}(x)|^{2\theta} \sim |\sqrt{\mathbf{a}^T\Sigma(l)\mathbf{a}} N(0,1)|^{2\theta} \sim (\mathbf{a}^T\Sigma(l)\mathbf{a})^\theta |N(0,1)|^{2\theta}.$$

We proceed by induction over the layers. For $l = 1$,

$$\begin{aligned}
\mathbb{E}\Big[|f_i^{(1)}(y) - f_i^{(1)}(x)|^{2\theta}\Big] &= C_\theta(\mathbf{a}^T\Sigma(1)\mathbf{a})^\theta \\
&= C_\theta(\sigma_\omega^2\|y\|^2_{\mathbb{R}^I} - 2\sigma_\omega^2\langle y, x\rangle_{\mathbb{R}^I} + \sigma_\omega^2\|x\|^2_{\mathbb{R}^I})^\theta \\
&= C_\theta(\sigma_\omega^2)^\theta(\|y\|^2_{\mathbb{R}^I} - 2\langle y, x\rangle_{\mathbb{R}^I} + \|x\|^2_{\mathbb{R}^I})^\theta \\
&= C_\theta(\sigma_\omega^2)^\theta\|y - x\|^{2\theta}_{\mathbb{R}^I},
\end{aligned}$$

where C_θ is the θ-th moment of the chi-square distribution with one degree of freedom. By hypothesis ϕ is Lipschitz.

$$\int |u - v|^{2\theta} q^{(l-1)}(\mathrm{d}u, \mathrm{d}v) \leq H^{(l-1)}\|y - x\|^{2\theta}_{\mathbb{R}^I}.$$

Then,

$$\begin{aligned}
|f_i^{(l)}(y) - f_i^{(l)}(x)|^{2\theta} &\sim |N(0,1)|^{2\theta}(\mathbf{a}^T\Sigma(l)\mathbf{a})^\theta \\
&= |N(0,1)|^{2\theta}\Big(\sigma_\omega^2\int[|\phi(u)|^2 - 2\phi(u)\phi(v) + |\phi(v)|^2]q^{(l-1)}(\mathrm{d}u, \mathrm{d}v)\Big)^\theta \\
&= |N(0,1)|^{2\theta}\Big(\sigma_\omega^2\int|\phi(u) - \phi(v)|^2 q^{(l-1)}(\mathrm{d}u, \mathrm{d}v)\Big)^\theta \\
&\leq |N(0,1)|^{2\theta}(\sigma_\omega^2 L_\phi^2)^\theta\Big(\int|u - v|^2 q^{(l-1)}(\mathrm{d}u, \mathrm{d}v)\Big)^\theta \\
&\leq |N(0,1)|^{2\theta}(\sigma_\omega^2 L_\phi^2)^\theta\int|u - v|^{2\theta} q^{(l-1)}(\mathrm{d}u, \mathrm{d}v) \\
&\leq |N(0,1)|^{2\theta}(\sigma_\omega^2 L_\phi^2)^\theta H^{(l-1)}\|y - x\|^{2\theta}_{\mathbb{R}^I}.
\end{aligned}$$

Thus we conclude

$$\mathbb{E}\Big[|f_i^{(l)}(y) - f_i^{(l)}(x)|^{2\theta}\Big] \leq H^{(l)}\|y - x\|^{2\theta}_{\mathbb{R}^I},$$

where the constant $H^{(l)}$ can be explicitly derived by solving the following system

$$\begin{cases} H^{(1)} = C_\theta(\sigma_\omega^2)^\theta \\ H^{(l)} = C_\theta(\sigma_\omega^2 L_\phi^2)^\theta H^{(l-1)}. \end{cases}$$

It is easy to get $H^{(l)} = C_\theta^l(\sigma_\omega^2)^{l\theta}(L_\phi^2)^{(l-1)\theta}$. Notice that this quantity does not depend on i. Therefore, by Proposition 2, by placing $\alpha = 2\theta$ and $\beta = 2\theta - I$, for every $\theta > I/2$ (β needs to be positive then we take $\theta > I/2$) there exists a continuous version $f_i^{(l)(\theta)}$ of the process $f_i^{(l)}$ with $\mathbb{P}$-a.s. locally γ-Hölder paths for every $0 < \gamma < 1 - \frac{I}{2\theta}$.

- Thus $f_i^{(l)(\theta)}$ and $f_i^{(l)}$ are indistinguishable (same trajectories), i.e there exists $\Omega^{(\theta)} \subset \Omega$ with $\mathbb{P}(\Omega^{(\theta)}) = 1$ such that for each $\omega \in \Omega^{(\theta)}$, $x \mapsto f_i^{(l)}(x)(\omega)$ is locally γ-Hölder for each $0 < \gamma < 1 - \frac{I}{2\theta}$.
- Define $\Omega^\star = \bigcap_{\theta > I/2} \Omega^{(\theta)}$, then for each $0 < \delta_0 < 1$ there exists θ_0 such that $\delta_0 < 1 - \frac{I}{2\theta_0} < 1$, thus for each $\omega \in \Omega^\star \subset \Omega^{(\theta_0)}$, the trajectory $x \mapsto f_i^{(l)}(x)(\omega)$ is locally δ_0-Hölder continuous.

By Proposition 2 we can conclude that $f_i^{(l)}$ has a continuous version and the latter is $\mathbb{P}$-a.s locally γ-Hölder continuous for every $0 < \gamma < 1$.

SM B.3

Fix $i \geq 1, l \geq 1$. We apply Proposition 3 to show the uniform tightness of the sequence $(f_i^{(l)}(n))_{n\geq 1}$ in $C(\mathbb{R}^I; \mathbb{R})$. By Lemma 2 $f_i^{(l)}(1), f_i^{(l)}(2), \ldots$ are random elements in $C(\mathbb{R}^I; \mathbb{R})$. First we show that the sequence $f(0_{\mathbb{R}^I}, n)_{n\geq 1}$ is uniformly tight in $\mathbb{R}$. We use the following statement from (Dudley, 2002, Theorem 11.5.3)

Proposition 4. *Let (C, ρ) be a metric space and suppose $f(n) \xrightarrow{d} f$ where $f(n)$ is tight for all n. Then $f(n)_{n\geq 1}$ is uniformly tight.*

Since $(\mathbb{R}, |\cdot|)$ is Polish every probability measure is tight, then $f(0_{\mathbb{R}^I}, n)$ is tight in $\mathbb{R}$ for every n. Moreover, by Lemma 1 $f_i(0_{\mathbb{R}^I}, n)_{n\geq 1} \xrightarrow{d} f_i^{(l)}(0_{\mathbb{R}^I})$, then by Proposition (4) $f(0_{\mathbb{R}^I}, n)_{n\geq 1}$ is uniformly tight in $\mathbb{R}$. In order to apply Proposition 3 it remains to show that there exist two values $\alpha > 0$ and $\beta > 0$, and a constant $H^{(l)} > 0$ such that

$$\mathbb{E}\Big[|f_i^{(l)}(y,n) - f_i^{(l)}(x,n)|^\alpha\Big] \leq H^{(l)} \|y - x\|_{\mathbb{R}^I}^{I+\beta}, \quad x, y \in \mathbb{R}^I, n \in \mathbb{N}$$

uniformly in n. The first idea could be try to bound (uniformly in n) the expected value of $H_i^{(l)}(n)$ obtained in (10), but this turns out to be very difficult. Thus we choose another way. Take two points $x, y \in \mathbb{R}^I$. From (8) we know that $f_i^{(l)}(y,n)|f_{1,\ldots,n}^{(l-1)} \sim N(0, \sigma_y^2(l,n))$ and $f_i^{(l)}(x,n)|f_{1,\ldots,n}^{(l-1)} \sim N(0, \sigma_x^2(l,n))$ with joint distribution $N_2(\mathbf{0}, \Sigma(l,n))$, where

$$\Sigma(1) = \begin{bmatrix} \sigma_x^2(1) & \Sigma(1)_{x,y} \\ \Sigma(1)_{x,y} & \sigma_y^2(1) \end{bmatrix},$$
$$\Sigma(l) = \begin{bmatrix} \sigma_x^2(l,n) & \Sigma(l,n)_{x,y} \\ \Sigma(l,n)_{x,y} & \sigma_y^2(l,n) \end{bmatrix},$$

with,

$$\begin{aligned}
\sigma_x^2(1) &= \sigma_b^2 + \sigma_\omega^2 \|x\|_{\mathbb{R}^I}^2, \\
\sigma_y^2(1) &= \sigma_b^2 + \sigma_\omega^2 \|y\|_{\mathbb{R}^I}^2, \\
\Sigma(1)_{x,y} &= \sigma_b^2 + \sigma_\omega^2 \langle x, y \rangle_{\mathbb{R}^I}, \\
\sigma_x^2(l,n) &= \sigma_b^2 + \frac{\sigma_\omega^2}{n} \sum_{j=1}^n |\phi \circ f_j^{(l-1)}(x,n)|^2, \\
\sigma_y^2(l,n) &= \sigma_b^2 + \frac{\sigma_\omega^2}{n} \sum_{j=1}^n |\phi \circ f_j^{(l-1)}(y,n)|^2, \\
\Sigma(l,n)_{x,y} &= \sigma_b^2 + \frac{\sigma_\omega^2}{n} \sum_{j=1}^n \phi(f_j^{(l-1)}(x,n)) \phi(f_j^{(l-1)}(y,n))
\end{aligned}$$

Defining $\mathbf{a}^T = [1, -1]$ we have that $f_i^{(l)}(y,n)|f_{1,\ldots,n}^{(l-1)} - f_i^{(l)}(x,n)|f_{1,\ldots,n}^{(l-1)}$ is distributed as $N(\mathbf{a}^T\mathbf{0}, \mathbf{a}^T\Sigma(l,n)\mathbf{a})$, where

$$\mathbf{a}^T \Sigma(l,n) \mathbf{a} = \sigma_y^2(l,n) - 2\Sigma(l,n)_{x,y} + \sigma_x^2(l,n).$$

Consider $\alpha = 2\theta$ with θ integer. Thus

$$\Big|f_i^{(l)}(y,n)|f_{1,\dots,n}^{(l-1)} - f_i^{(l)}(x,n)|f_{1,\dots,n}^{(l-1)}\Big|^{2\theta} \sim |\sqrt{\mathbf{a}^T\Sigma(l,n)\mathbf{a}}N(0,1)|^{2\theta} \sim (\mathbf{a}^T\Sigma(l,n)\mathbf{a})^{\theta}|N(0,1)|^{2\theta}.$$

Start first with the case $l = 1$.

$$\begin{aligned}\mathbb{E}\Big[|f_i^{(1)}(y,n) - f_i^{(1)}(x,n)|^{2\theta}\Big] &= C_\theta(\mathbf{a}^T\Sigma(1)\mathbf{a})^{\theta} \\ &= C_\theta(\sigma_\omega^2\|y\|^2_{\mathbb{R}^I} - 2\sigma_\omega^2\langle y,x\rangle_{\mathbb{R}^I} + \sigma_\omega^2\|x\|^2_{\mathbb{R}^I})^{\theta} \\ &= C_\theta(\sigma_\omega^2)^{\theta}(\|y\|^2_{\mathbb{R}^I} - 2\langle y,x\rangle_{\mathbb{R}^I} + \|x\|^2_{\mathbb{R}^I})^{\theta} \\ &= C_\theta(\sigma_\omega^2)^{\theta}\|y-x\|^{2\theta}_{\mathbb{R}^I},\end{aligned}$$

where C_θ is the θ-th moment of the chi-square distribution with one degree of freedom. Set $H^{(1)} = C_\theta(\sigma_\omega^2)^{\theta}$. By hypothesis induction suppose that for every $j \geq 1$

$$\mathbb{E}\Big[|f_j^{(l-1)}(y,n) - f_j^{(l-1)}(x,n)|^{2\theta}\Big] \leq H^{(l-1)}\|y-x\|^{2\theta}_{\mathbb{R}^I}.$$

By hypothesis ϕ is Lipschitz, then

$$\begin{aligned}\mathbb{E}\Big[|f_i^{(l)}(y,n) - f_i^{(l)}(x,n)|^{2\theta}\Big|f_{1,\dots,n}^{(l-1)}\Big] &= C_\theta(\mathbf{a}^T\Sigma(l,n)\mathbf{a})^{\theta} \\ &= C_\theta\Big(\sigma_y^2(l,n) - 2\Sigma(l,n)_{x,y} + \sigma_x^2(l,n)\Big)^{\theta} \\ &= C_\theta\Big(\frac{\sigma_\omega^2}{n}\sum_{j=1}^{n}\Big|\phi\circ f_j^{(l-1)}(y,n) - \phi\circ f_j^{(l-1)}(x,n)\Big|^2\Big)^{\theta} \\ &\leq C_\theta\Big(\frac{\sigma_\omega^2 L_\phi^2}{n}\sum_{j=1}^{n}\Big|f_j^{(l-1)}(y,n) - f_j^{(l-1)}(x,n)\Big|^2\Big)^{\theta} \\ &= C_\theta\frac{(\sigma_\omega^2 L_\phi^2)^{\theta}}{n^{\theta}}\Big(\sum_{j=1}^{n}\Big|f_j^{(l-1)}(y,n) - f_j^{(l-1)}(x,n)\Big|^2\Big)^{\theta} \\ &\leq C_\theta\frac{(\sigma_\omega^2 L_\phi^2)^{\theta}}{n}\sum_{j=1}^{n}\Big|f_j^{(l-1)}(y,n) - f_j^{(l-1)}(x,n)\Big|^{2\theta}.\end{aligned}$$

Using the induction hypothesis

$$\begin{aligned}\mathbb{E}\Big[|f_i^{(l)}(y,n) - f_i^{(l)}(x,n)|^{2\theta}\Big] &= \mathbb{E}\Big[\mathbb{E}\Big[|f_i^{(l)}(y,n) - f_i^{(l)}(x,n)|^{2\theta}\Big|f_{1,\dots,n}^{(l-1)}\Big]\Big] \\ &\leq C_\theta\frac{(\sigma_\omega^2 L_\phi^2)^{\theta}}{n}\sum_{j=1}^{n}\mathbb{E}\Big[|f_j^{(l-1)}(y,n) - f_j^{(l-1)}(x,n)|^{2\theta}\Big] \\ &\leq C_\theta(\sigma_\omega^2 L_\phi^2)^{\theta}H^{(l-1)}\|y-x\|^{2\theta}_{\mathbb{R}^I}.\end{aligned}$$

We can get the constant $H^{(l)}$ by solving the same system as (12), obtaining $H^{(l)} = C_\theta^l(\sigma_\omega^2)^{l\theta}(L_\phi^2)^{(l-1)\theta}$ which does not depend on n. By Proposition 3 setting $\alpha = 2\theta$ and $\beta = 2\theta - I$, since β must be a positive constant, it is sufficient to take $\theta > I/2$ and this concludes the proof.

SM C

Fix k inputs $\mathbf{X} = [x^{(1)}, \dots, x^{(k)}]$ and a layer l. We show that as $n \to +\infty$

$$\Big(f_i^{(l)}(\mathbf{X},n)\Big)_{i\geq 1} \xrightarrow{d} \bigotimes_{i=1}^{\infty} N_k(\mathbf{0},\Sigma(l))$$

where $\bigotimes$ denotes the product measure and with $\Sigma(l)$ as in (7). We prove this statement by proving the n large asymptotic behaviour of any finite linear combination of the $f_i^{(l)}(\mathbf{X},n)$'s, for $i \in \mathcal{L} \subset \mathbb{N}$.

See, e.g. Billingsley (1999) for details. Following the notation of Matthews et al. (2018b), consider a finite linear combination of the function values without the bias, i.e.,

$$\mathcal{T}^{(l)}(\mathcal{L}, p, \mathbf{X}, n) = \sum_{i \in \mathcal{L}} p_i [f_i^{(l)}(\mathbf{X}, n) - b_i^{(l)} \mathbf{1}].$$

Then for the first layer we write

$$\begin{aligned}
\mathcal{T}^{(1)}(\mathcal{L}, p, \mathbf{X}) &= \sum_{i \in \mathcal{L}} p_i \Big[\sum_{j=1}^{I} \omega_{i,j}^{(1)} \mathbf{x}_j \Big] \\
&= \sum_{j=1}^{I} \gamma_j^{(1)}(\mathcal{L}, p, \mathbf{X}),
\end{aligned}$$

where

$$\gamma_j^{(1)}(\mathcal{L}, p, \mathbf{X}) = \sum_{i \in \mathcal{L}} p_i \omega_{i,j}^{(l)} \mathbf{x}_j.$$

and for any $l \geq 2$

$$\begin{aligned}
\mathcal{T}^{(l)}(\mathcal{L}, p, \mathbf{X}, n) &= \sum_{i \in \mathcal{L}} p_i \Big[\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \omega_{i,j}^{(l)} (\phi \bullet f_j^{(l-1)}(\mathbf{X}, n)) \Big] \\
&= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \gamma_j^{(l)}(\mathcal{L}, p, \mathbf{X}, n),
\end{aligned}$$

where

$$\gamma_j^{(l)}(\mathcal{L}, p, \mathbf{X}, n) = \sum_{i \in \mathcal{L}} p_i \omega_{i,j}^{(l)} (\phi \bullet f_j^{(l-1)}(\mathbf{X}, n)).$$

For the first layer we get

$$\begin{aligned}
\varphi_{\mathcal{T}^{(1)}(\mathcal{L}, p, \mathbf{X})}(\mathbf{t}) &= \mathbb{E}\Big[e^{\mathbf{i}\mathbf{t}^T \mathcal{T}^{(1)}(\mathcal{L}, p, \mathbf{X})} \Big] \\
&= \mathbb{E}\Big[\exp\Big\{ \mathbf{i}\mathbf{t}^T \Big[\sum_{j=1}^{I} \sum_{i \in \mathcal{L}} p_i \omega_{i,j}^{(1)} \mathbf{x}_j \Big] \Big\} \Big] \\
&= \prod_{j=1}^{I} \prod_{i \in \mathcal{L}} \mathbb{E}\Big[\exp\Big\{ \mathbf{i}\mathbf{t}^T \Big[p_i \omega_{i,j}^{(1)} \mathbf{x}_j \Big] \Big\} \Big] \\
&= \prod_{j=1}^{I} \prod_{i \in \mathcal{L}} \exp\Big\{ -\frac{\sigma_\omega^2}{2} p_i^2 \Big(\mathbf{t}^T \mathbf{x}_j \Big)^2 \Big\} \\
&= \exp\Big\{ -\frac{\sigma_\omega^2}{2n} \sum_{i \in \mathcal{L}} p_i^2 \sum_{j=1}^{n} \Big(\mathbf{t}^T \mathbf{x}_j \Big)^2 \Big\} \\
&= \exp\Big\{ -\frac{1}{2} \mathbf{t}^T \Theta(\mathcal{L}, p, 1) \mathbf{t} \Big\},
\end{aligned}$$

i.e.

$$\mathcal{T}^{(1)}(\mathcal{L}, p, \mathbf{X}) \stackrel{d}{=} N_k(\mathbf{0}, \Theta(\mathcal{L}, p, 1)),$$

with $k \times k$ covariance matrix with element in the i-th row and j-th column as follows

$$\Theta_{i,j}(\mathcal{L}, p, 1) = p^T p \sigma_\omega^2 \langle x^{(i)}, x^{(j)} \rangle_{\mathbb{R}^I},$$

where $p^T p = \sum_{i \in \mathcal{L}} p_i^2$. For $l \geq 2$ we get

$$\begin{aligned}
\varphi_{\mathcal{T}^{(l)}(\mathcal{L},p,\mathbf{X},n)|f_{1,\dots,n}^{(l-1)}}(\mathbf{t}) &= \mathbb{E}\Big[e^{\mathrm{i}\mathbf{t}^T \mathcal{T}^{(l)}(\mathcal{L},p,\mathbf{X},n)}|f_{1,\dots,n}^{(l-1)}\Big] \\
&= \mathbb{E}\Big[\exp\Big\{\mathrm{i}\mathbf{t}^T\Big[\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\sum_{i\in\mathcal{L}} p_i \omega_{i,j}^{(l)}(\phi \bullet f_j^{(l-1)}(\mathbf{X},n))\Big]\Big\}|f_{1,\dots,n}^{(l-1)}\Big] \\
&= \prod_{j=1}^{n}\prod_{i\in\mathcal{L}} \mathbb{E}\Big[\exp\Big\{\mathrm{i}\mathbf{t}^T\Big[\frac{1}{\sqrt{n}} p_i \omega_{i,j}^{(l)}(\phi \bullet f_j^{(l-1)}(\mathbf{X},n))\Big]\Big\}|f_{1,\dots,n}^{(l-1)}\Big] \\
&= \prod_{j=1}^{n}\prod_{i\in\mathcal{L}} \exp\Big\{-\frac{\sigma_\omega^2}{2n} p_i^2 \Big(\mathbf{t}^T(\phi \bullet f_j^{(l-1)}(\mathbf{X},n))\Big)^2\Big\} \\
&= \exp\Big\{-\frac{\sigma_\omega^2}{2n}\sum_{i\in\mathcal{L}} p_i^2 \sum_{j=1}^{n}\Big(\mathbf{t}^T(\phi \bullet f_j^{(l-1)}(\mathbf{X},n))\Big)^2\Big\} \\
&= \exp\Big\{-\frac{1}{2}\mathbf{t}^T \Theta(\mathcal{L},p,l,n)\mathbf{t}\Big\},
\end{aligned}$$

i.e.

$$\mathcal{T}^{(l)}(\mathcal{L},p,\mathbf{X},n)|f_{1,\dots,n}^{(l-1)} \overset{d}{=} N_k(\mathbf{0}, \Theta(\mathcal{L},p,l,n)),$$

with $k \times k$ covariance matrix with element in the i-th row and j-th column as follows

$$\Theta_{i,j}(\mathcal{L},p,l,n) = p^T p \frac{\sigma_\omega^2}{n}\Big\langle(\phi \bullet \mathbf{F}_i^{(l-1)}(\mathbf{X},n)), (\phi \bullet \mathbf{F}_j^{(l-1)}(\mathbf{X},n))\Big\rangle_{\mathbb{R}^n},$$

where $p^T p = \sum_{i \in \mathcal{L}} p_i^2$. Thus, along lines similar to the proof of the large n asymptotics for the $i-th$ coordinate (just replacing $\sigma_b^2 \leftarrow 0$ and $\sigma_\omega^2 \leftarrow p^T p \sigma_\omega^2$), we have that for any $l \geq 2$, as $n \to +\infty$,

$$\begin{aligned}
\varphi_{\mathcal{T}^{(l)}(\mathcal{L},p,\mathbf{X},n)}(\mathbf{t}) &\to \exp\Big\{-\frac{1}{2} p^T p \sigma_\omega^2 \int \Big(\mathbf{t}^T(\phi \bullet f)\Big)^2 q^{(l-1)}(\mathrm{d}f)\Big]\Big\} \\
&= \exp\Big\{-\frac{1}{2}\mathbf{t}^T \Theta(\mathcal{L},p,l)\mathbf{t}\Big\},
\end{aligned}$$

i.e. $\mathcal{T}^{(l)}(\mathcal{L},p,\mathbf{X},n)$ converges weakly to a k-dimensional Gaussian distribution with mean $\mathbf{0}$ and $k \times k$ covariance matrix $\Theta(\mathcal{L},p,l)$ with elements

$$\Theta_{i,j}(\mathcal{L},p,l) = p^T p \sigma_\omega^2 \int \phi(f_i)\phi(f_j) q^{(l-1)}(\mathrm{d}f),$$

where $q^{(l-1)}(\mathrm{d}f) = q^{(l-1)}(\mathrm{d}f_1, \dots, \mathrm{d}f_k) = N_k(\mathbf{0}, \Theta(\mathcal{L},p,l-1))\mathrm{d}f$. To complete the proof just observe that $\Theta(\mathcal{L},p,l) = p^T p \Sigma(l)$.

SM D.1

We will use, without explicit mention, that the series $\sum_{i=1}^{\infty} q^i$ converges when $|q| < 1$. In particular when $q = 1/2$ the series sum to 1. Fix, $l \geq 1$ and $n \in \mathbb{N}$. We prove that there exists a random variable $H^{(l)}(n)$ such that

$$d\big(\mathbf{F}^{(l)}(x,n), \mathbf{F}^{(l)}(y,n)\big)_\infty \leq H^{(l)}(n)\|x-y\|_{\mathbb{R}^I}, \quad x,y \in \mathbb{R}^I, \mathbb{P}-a.s.$$

It immediately derives from the Lipschitzianity of each component, indeed by (9) we get

$$\begin{aligned}
d\big(\mathbf{F}^{(l)}(x,n), \mathbf{F}^{(l)}(y,n)\big)_\infty &= \sum_{i=1}^{\infty}\frac{1}{2^i}\frac{|f_i^{(l)}(x,n) - f_i^{(l)}(y,n)|}{1 + |f_i^{(l)}(x,n) - f_i^{(l)}(y,n)|} \\
&\leq \sum_{i=1}^{\infty}\frac{1}{2^i}|f_i^{(l)}(x,n) - f_i^{(l)}(y,n)| \\
&\leq \|x-y\|_{\mathbb{R}^I}\sum_{i=1}^{\infty}\frac{1}{2^i} H_i^{(l)}(n).
\end{aligned}$$

It remains to show that the series $\sum_{i=1}^{\infty} \frac{1}{2^i} H_i^{(l)}(n)$ converges almost surely. By (10) we get

$$\begin{aligned}\sum_{i=1}^{\infty} \frac{1}{2^i} H_i^{(l)}(n) &= \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{L_\phi}{\sqrt{n}} \sum_{j=1}^{n} |\omega_{i,j}^{(l)}| H_j^{(l-1)}(n) \\ &= \frac{L_\phi}{\sqrt{n}} \sum_{j=1}^{n} H_j^{(l-1)}(n) \sum_{i=1}^{\infty} \frac{|\omega_{i,j}^{(l)}|}{2^i}.\end{aligned}$$

It remains to show the convergence almost surely of the series $\sum_{i=1}^{\infty} \frac{|\omega_{i,j}^{(l)}|}{2^i}$. We apply the three-series Kolmogorov criterion (Kallenberg, 2002, Theorem 4.18). Call $X_i := \frac{|\omega_{i,j}^{(l)}|}{2^i}$

- By Markov inequality $\mathbb{P}(X_i > 1) \leq \mathbb{E}[X_i] = \frac{\mathbb{E}[|N(0,\sigma_\omega^2)|]}{2^i}$, thus $\sum_{i=1}^{\infty} \mathbb{P}(X_i > 1) \leq \mathbb{E}[|N(0,\sigma_\omega^2)|] < \infty$
- Call $Y_i = X_i \mathbb{I}_{\{X_i \leq 1\}} \leq X_i$. Then $\sum_{i=1}^{\infty} \mathbb{E}[Y_i] \leq \sum_{i=1}^{\infty} \mathbb{E}[X_i] = \sum_{i=1}^{\infty} \frac{\mathbb{E}[|N(0,\sigma_\omega)|]}{2^i} = \mathbb{E}[|N(0,\sigma_\omega^2)|] < \infty$
- $\mathrm{V}(Y_i) = \mathbb{E}[Y_i^2] - \mathbb{E}^2[Y_i]$, thus $\sum_{i=1}^{\infty} \mathrm{V}(Y_i) = \sum_{i=1}^{\infty} \mathbb{E}[Y_i^2] - \sum_{i=1}^{\infty} \mathbb{E}^2[Y_i]$. The first series converges since $\mathbb{E}[Y_i^2] \leq \mathbb{E}[X_i^2] = \frac{\sigma_\omega^2 \mathbb{E}[\chi^2(1)]}{4^i} = \frac{\sigma_\omega^2}{4^i}$ (then $\sum \mathbb{E} Y_i \leq \sigma_\omega^2 \sum \frac{1}{4^i} < \infty$), and the other series converges since $0 < \mathbb{E}[Y_i] \leq \mathbb{E}[X_i]$ implies $\mathbb{E}^2[Y_i] \leq \mathbb{E}^2[X_i] = \frac{\mathbb{E}^2[|N(0,\sigma_\omega^2)|]}{4^i}$ (then $\sum \mathbb{E}^2[Y_i] \leq \mathbb{E}^2[|N(0,\sigma_\omega^2)|] \sum \frac{1}{4^i} < \infty$).

Denoting $Q_j^{(l)} = \sum_{i=1}^{\infty} \frac{|\omega_{i,j}^{(l)}|}{2^i}$ and by setting $H^{(l)}(n) := \frac{L_\phi}{\sqrt{n}} \sum_{j=1}^{n} H_j^{(l-1)}(n) Q_j^{(l)}$ we complete the proof.

SM D.2

Fix $l \geq 1$. We show the continuity of the limiting process $\mathbf{F}^{(l)}$ by applying Proposition 2. We will use, without explicit mention, that the function $r \mapsto \frac{r}{1+r}$ is bounded by 1 for $r > 0$. Take two inputs $x, y \in \mathbb{R}^I$ and fix $\alpha \geq 12$ even integer. Since $\sum_{i=1}^{\infty} \frac{1}{2^i} \frac{|f^{(l)}(x) - f_i^{(l)}(y)|}{1+|f^{(l)}(x) - f_i^{(l)}(y)|} < \sum_{i=1}^{\infty} \frac{1}{2^i} = 1$ and, by Jensen inequality, also $\sum_{i=1}^{\infty} \frac{1}{2^i} \big(\frac{|f^{(l)}(x) - f_i^{(l)}(y)|}{1+|f^{(l)}(x) - f_i^{(l)}(y)|}\big)^\alpha < \sum_{i=1}^{\infty} \frac{1}{2^i} = 1$, we get

$$\begin{aligned}d\big(\mathbf{F}^{(l)}(x), \mathbf{F}^{(l)}(y)\big)_\infty^\alpha &= \Big(\sum_{i=1}^{\infty} \frac{1}{2^i} \frac{|f_i^{(l)}(x) - f_i^{(l)}(y)|}{1 + |f_i^{(l)}(x) - f_i^{(l)}(y)|}\Big)^\alpha \\ &\leq \sum_{i=1}^{\infty} \frac{1}{2^i} \Big(\frac{|f_i^{(l)}(x) - f_i^{(l)}(y)|}{1 + |f_i^{(l)}(x) - f_i^{(l)}(y)|}\Big)^\alpha \\ &\leq \sum_{i=1}^{\infty} \frac{1}{2^i} |f_i^{(l)}(x) - f_i^{(l)}(y)|^\alpha\end{aligned}$$

Thus, by applying monotone convergence theorem to the positive increasing sequence $g(N) = \sum_{i=1}^{N} \frac{1}{2^i} |f_i^{(l)}(x) - f_i^{(l)}(y)|^\alpha$ (which allows to exchange $\mathbb{E}$ and $\sum_{i=1}^{\infty}$), we get

$$
\begin{aligned}
\mathbb{E}\Big[d\big(\mathbf{F}^{(l)}(x), \mathbf{F}^{(l)}(y)\big)_{\infty}^{\alpha}\Big] &\le \mathbb{E}\Big[\sum_{i=1}^{\infty} \frac{1}{2^i} |f^{(l)}(x) - f_i^{(l)}(y)|^{\alpha}\Big] \\
&= \mathbb{E}\Big[\lim_{N\to\infty} \sum_{i=1}^{N} \frac{1}{2^i} |f_i^{(l)}(x) - f_i^{(l)}(y)|^{\alpha}\Big] \\
&= \lim_{N\to\infty} \mathbb{E}\Big[\sum_{i=1}^{N} \frac{1}{2^i} |f_i^{(l)}(x) - f_i^{(l)}(y)|^{\alpha}\Big] \\
&= \sum_{i=1}^{\infty} \frac{1}{2^i} \mathbb{E}\Big[|f_i^{(l)}(x) - f_i^{(l)}(y)|^{\alpha}\Big] \\
&= \sum_{i=1}^{\infty} \frac{1}{2^i} H^{(l)} \|x-y\|_{\mathbb{R}^I}^{\alpha} \\
&= H^{(l)} \|x-y\|_{\mathbb{R}^I}^{\alpha}
\end{aligned}
$$

where we used (11) and the fact that $H^{(l)}$ does not depend on i (see (12)).

Therefore, by Proposition 2, for each $\alpha > I$, setting $\beta = \alpha - I$ (since β needs to be positive, it is sufficient to choose $\alpha > I$) $\mathbf{F}^{(l)}$ has a continuous version $\mathbf{F}^{(l)(\theta)}$ and the latter is $\mathbb{P}$-a.s locally γ-Hölder continuous for every $0 < \gamma < 1 - \frac{I}{\alpha}$.

- Thus $\mathbf{F}^{(l)(\alpha)}$ and $\mathbf{F}^{(l)}$ are indistinguishable (same trajectories), i.e there exists $\Omega^{(\alpha)} \subset \Omega$ with $\mathbb{P}(\Omega^{(\alpha)}) = 1$ such that for each $\omega \in \Omega^{(\alpha)}$, $x \mapsto \mathbf{F}^{(l)}(x)(\omega)$ is locally γ-Hölder for each $0 < \gamma < 1 - \frac{I}{\alpha}$.
- Define $\Omega^{\star} = \bigcap_{\alpha > I} \Omega^{(\alpha)}$, then for each $0 < \delta_0 < 1$ there exists α_0 such that $\delta_0 < 1 - \frac{I}{\alpha_0} < 1$, thus for each $\omega \in \Omega^{\star} \subset \Omega^{(\alpha_0)}$, the trajectory $x \mapsto \mathbf{F}^{(l)}(x)(\omega)$ is locally δ_0-Hölder continuous.

By Proposition 2 we can conclude that $\mathbf{F}^{(l)}$ has a continuous version and the latter is $\mathbb{P}$-a.s locally γ-Hölder continuous for every $0 < \gamma < 1$.

SM E

General introduction to Daniell-Kolmogorov extension theorem

Let X be a set of indexes and $\{(E_x, \mathcal{E}_x)\}_{x\in X}$ measurable spaces. On $E := \times_{x\in X} E_x$ we can consider the σ-algebra $\mathcal{E} := \bigotimes_{x\in X} \mathcal{E}_x$ that is

$$
\mathcal{E} = \sigma(\pi_x, x \in X) = \sigma\Big(\bigcup_{x\in X} \pi_x^{-1}(\mathcal{E}_x)\Big)
$$

where for each $x \in X$, $\pi_x : E \to E_x, \omega := (\omega_x)_{x\in X} \mapsto \pi_x(\omega) = \omega_x$. $\mathcal{E}$ is generated by measurable rectangles. A measurable rectangle A is of the form

$$
A := \times_{x\in X} A_x \text{ such that only a finite number of } A_x \in \mathcal{E}_x \text{ are different from } E_x
$$

σ-algebra on the space of functions

Fix $X = \mathbb{R}^I$ and (S, d) Polish space. We consider the measurable sets $\{(E_x, \mathcal{E}_x)\}_{x\in X} = \{(S, \mathcal{B}(S))\}_{x\in\mathbb{R}^I}$ thus we can construct a measurable space

$$
(E, \mathcal{E}) = (\times_{x\in X} E_x, \bigotimes_{x\in X} \mathcal{E}_x) = \Big(S^{\mathbb{R}^I}, \mathcal{B}(S^{\mathbb{R}^I})\Big)
$$

where $S^{\mathbb{R}^I} = \times_{x\in\mathbb{R}^I} S$ is the set of all functions from $\mathbb{R}^I$ into S and

$$\begin{aligned}\mathcal{B}(S^{\mathbb{R}^I}) &:= \bigotimes_{x\in\mathbb{R}^I} \mathcal{B}(S)\\ &= \sigma\Big(\bigcup_{x\in\mathbb{R}^I} \pi_x^{-1}(\mathcal{B}(S))\Big)\\ &= \sigma\Big(\Big\{A := \times_{x\in X} A_x \text{ such that only a finite number of } A_x \text{ are different from } S\ \Big\}\Big)\end{aligned}$$

An example of measurable rectangle is

$$A = S \times A_{x^{(1)}} \times S \times A_{x^{(2)}} \times S \times S \times \cdots \times A_{x^{(k)}} \times S \times S \times \dots$$

where $k \in \mathbb{N}$ and only for $x^{(1)}, \dots, x^{(k)}$ the cartesian products are different to S.

Denote by $Z = (Z_x)_{x\in\mathbb{R}^I}$, $Z_x : (\Omega, \mathcal{H}, \mathbb{P}) \to S$ any stochastic process of interest, such as $f_i^{(l)}(n)$ or $f_i^{(l)}$ for some $l \geq 1$, $i \geq 1$ and $n \geq 1$ when $S = (\mathbb{R}, |\cdot|)$, or even $\mathbf{F}^{(l)}(n)$ or $\mathbf{F}^{(l)}$ for $l \geq 1$ and $n \geq 1$ when $S = (\mathbb{R}^\infty, \|\cdot\|_\infty)$. Consider the finite-dimensional distributions of Z

$$\Lambda = \{P^Z_{x^{(1)},\dots,x^{(k)}} \text{ on } \mathcal{B}(S^k) | x^{(j)} \in \mathbb{R}^I, j \in \{1, \dots, k\}, k \in \mathbb{N}\}$$

If Λ is consistent in the sense of Kolmogorov theorem, then there exists an unique probability measure $\mathbb{P}'$ on $(S^{\mathbb{R}^I}, \mathcal{B}(S^{\mathbb{R}^I}))$ such that the canonical process $Z' = (Z'_x)_{x\in\mathbb{R}^I}$, $Z'_x : S^{\mathbb{R}^I} \to S, \omega \mapsto Z'_x(\omega) = \omega(x)$ on $(S^{\mathbb{R}^I}, \mathcal{B}(S^{\mathbb{R}^I}), \mathbb{P}')$ has finite-dimensional distributions that coincide with Λ.

SM E.1 : Existence of a probability measure on $S^{\mathbb{R}^I}$ for the sequence processes

Fix $S = \mathbb{R}$. Fix a layer l, a unit $i \geq 1$ on that layer and $n \in \mathbb{N}$. We want to prove that there exists a probability measure $\mathbb{P}^{(i,l,n)}$ on $(\mathbb{R}^{\mathbb{R}^I}, \mathcal{B}(\mathbb{R}^{\mathbb{R}^I}))$ such that the associated canonical process $\Theta_x^{(i,l,n)} : \mathbb{R}^{\mathbb{R}^I} \to \mathbb{R}$, $\omega \mapsto \omega(x)$ has finite-dimensional distributions that coincide with

$$\Lambda^{(i,l,n)} = \Big\{P^{(i,l,n)}_{x^{(1)},\dots,x^{(k)}}\Big\}_{k\in\mathbb{N}},$$

where $P^{(i,l,n)}_{x^{(1)},\dots,x^{(k)}}$ is the distribution of $f_i^{(l)}(\mathbf{X}, n)$. We do not know the exact form of this distribution but we know the distribution of the conditioned random variable $f_i^{(l)}(\mathbf{X}, n)|f^{(l-1)}_{1,\dots,n}$ (see (8)). Thus, since from (8) the distribution of $f_i^{(1)}(\mathbf{X})$ is well known, proceeding by induction it is sufficient to prove the existence of two probability measures $\mathbb{P}^{(i,1,n)}$ and $\mathbb{P}^{(i,l,n)}_{|l-1}$ on $(\mathbb{R}^{\mathbb{R}^I}, \mathcal{B}(\mathbb{R}^{\mathbb{R}^I}))$ such that the associated canonical processes $\Theta_x^{(i,1,n)}$, and $\Theta_x^{(i,1,n)|l-1}$ have finite-dimensional distributions that coincide respectively with

$$\Lambda^{(i,1,n)} := \Big\{P^{(i,1,n)}_{x^{(1)},\dots,x^{(k)}}\Big\}_{k\in\mathbb{N}} \quad \text{and} \quad \Lambda^{(i,l,n)}_{|l-1} := \Big\{P^{(i,l,n,)|l-1}_{x^{(1)},\dots,x^{(k)}}\Big\}_{k\in\mathbb{N}},$$

where $P^{(i,1,n)}_{x^{(1)},\dots,x^{(k)}} = N_k(\mathbf{0}, \Sigma(1, \mathbf{X}))$ and $P^{(i,l,n)|l-1}_{x^{(1)},\dots,x^{(k)}} = N_k(\mathbf{0}, \Sigma(l, n, \mathbf{X}))$ defined on $\mathcal{B}(\mathbb{R}^k)$. Observe that, for simplicity of notation, we have always avoided to write the dependence of the covariance matrix on the inputs matrix $\mathbf{X}$, but in this case it is important to emphasize this. For the proof we defer to the limit case in the next subsection since the proof is the same step by step. When $S = \mathbb{R}^\infty$, recall that given a sequence of probability spaces $\{(\mathbb{R}^{\mathbb{R}^I}, \mathcal{B}(\mathbb{R}^{\mathbb{R}^I}), \mathbb{P}^{(i,l,n)})\}_{i\geq 1}$ there exists a unique probability measure $\mathbb{P}^{(l,n)}$ on $(\times_{i=1}^\infty \mathbb{R}^{\mathbb{R}^I}, \bigotimes_{i=1}^\infty \mathcal{B}(\mathbb{R}^{\mathbb{R}^I})) = \big((\mathbb{R}^\infty)^{\mathbb{R}^I}, \mathcal{B}((\mathbb{R}^\infty)^{\mathbb{R}^I})\big)$ such that, for each measurable rectangle $A = \times_{i=1}^\infty A_i$ where only for a finite number of i the set A_i is different from $\mathbb{R}^{\mathbb{R}^I}$, then $\mathbb{P}^{(l,n)}(A) = \prod_{i=1}^\infty \mathbb{P}^{(i,l,n)}(A_i)$. Moreover this probability is denoted as $\mathbb{P}^{(l,n)} =: \bigotimes_{i=1}^\infty \mathbb{P}^{(i,l,n)}$. This means that the existence of the stochastic processes $f_i^{(l)}(n)$ implies the existence of the stochastic processes $\mathbf{F}^{(l)}(n)$.

SM E.2 : Existence of a probability measure on $S^{\mathbb{R}^I}$ for the limit process

Note that, as observed in previous section, the existence of the stochastic processes $f_i^{(l)}$ on $(\mathbb{R}^{\mathbb{R}^I}, \mathcal{B}(\mathbb{R}^{\mathbb{R}^I}))$ implies the existence of the stochastic processes $\mathbf{F}^{(l)}$ on $\big((\mathbb{R}^\infty)^{\mathbb{R}^I}, \mathcal{B}((\mathbb{R}^\infty)^{\mathbb{R}^I})\big)$. Then we focus on the proof when $S = \mathbb{R}$. Fix a layer l and a unit $i \geq 1$ on that layer. We want to prove that there exists a probability measure $\mathbb{P}^{(i,l)}$ on $(\mathbb{R}^{\mathbb{R}^I}, \mathcal{B}(\mathbb{R}^{\mathbb{R}^I}))$ such that the canonical process $\Theta_x^{(i,l)} : \mathbb{R}^{\mathbb{R}^I} \to \mathbb{R}$, $\omega \mapsto \omega(x)$ has finite-dimensional distributions that coincide with

$$\Lambda^{(i,l)} = \Big\{ P^{(i,l)}_{x^{(1)},\dots,x^{(k)}} \Big\}_{k\in\mathbb{N}},$$

where $P^{(i,l)}_{x^{(1)},\dots,x^{(k)}}$ are the finite-dimensional distributions of $f_i^{(l)}$ determined in (7), i.e. $P^{(i,l)}_{x^{(1)},\dots,x^{(k)}} = N_k(\mathbf{0}, \Sigma(l, \mathbf{X}))$ defined on $\mathcal{B}(\mathbb{R}^k)$. By Daniell-Kolmogorov existence result (Kallenberg, 2002, Theorem 6.16) it is sufficient to prove that for each $k \in \mathbb{N}$ and for each $x^{(1)}, \dots, x^{(k)}$ elements on $\mathbb{R}^I$, then

$$\begin{aligned} &P^{(i,l)}_{x^{(1)},\dots,x^{(z)},\dots,x^{(k)}}(B^{(1)} \times \dots \times B^{(z-1)} \times \mathbb{R} \times B^{(z+1)} \times \dots \times B^{(k)}) \\ &= P^{(i,l)}_{x^{(1)},\dots,x^{(z-1)},x^{(z+1)},\dots,x^{(k)}}(B^{(1)} \times \dots \times B^{(z-1)} \times B^{(z+1)} \times \dots \times B^{(k)}), \end{aligned} \tag{18}$$

for every $z \in \{1, \dots, k\}$ and for every $B^{(j)} \in \mathcal{B}(\mathbb{R})$ for all $j = 1, \dots, k$, $j \neq z$. Fix $k \in \mathbb{N}$, k inputs $x^{(1)}, \dots, x^{(k)}$, $z \in \{1, \dots, k\}$ and $B^{(j)} \in \mathcal{B}(\mathbb{R})$ for all $j = 1, \dots, k$, $j \neq z$. Define the projection $\pi_{[z]} : \mathbb{R}^k \to \mathbb{R}^{k-1}$ such that $\pi_{[k]}(y_1, \dots, y_k) = [y_1, \dots, y_{z-1}, y_{z+1}, \dots y_k]^T$. Thus, condition (18) is equivalent to the following:

$$P^{(i,l)}_{x^{(1)},\dots,x^{(k)}} \circ \pi_{[z]} = P^{(i,l)}_{\pi_{[z]}(x^{(1)},\dots,x^{(k)})},$$

where on the left we have the image measure of $P^{(i,l)}_{x^{(1)},\dots,x^{(k)}}$ under $\pi_{[z]}$. We prove this by proving that the respective Fourier transformations coincide. In the following calculations we define $\mathbf{y} = [y_1, \dots, y_k]^T$, $\mathbf{y}_{[z]} = [y_1, \dots, y_{z-1}, y_{z+1}, \dots, y_k]^T$ and $\mathbf{t} = [t_1, \dots, t_k]^T$, $\mathbf{t}_{[k]} = [t_1, \dots, t_{z-1}, t_{z+1}, \dots, t_k]^T$, then by definition of image measure we get

$$\begin{aligned} \varphi_{\left(P^{(i,l)}_{x^{(1)},\dots,x^{(k)}} \circ \pi_{[z]}\right)}(\mathbf{t}_{[z]}) &= \int_{\mathbb{R}^{k-1}} e^{\mathrm{i}\mathbf{t}_{[z]}^T \mathbf{y}_{[z]}} \big(P^{(i,l)}_{x^{(1)},\dots,x^{(k)}} \circ \pi_{[z]}\big)(\mathrm{d}\mathbf{y}_{[z]}) \\ &= \int_{\mathbb{R}^k} e^{\mathrm{i}\mathbf{t}_{[z]}^T \pi_{[z]}(\mathbf{y})} P^{(i,l)}_{x^{(1)},\dots,x^{(k)}}(\mathrm{d}\mathbf{y}). \end{aligned}$$

Now, recalling that $\mathbf{1}_j$ is the $k \times 1$ vector with 1 in the j-th position and 0 otherwise, since $\pi_{[z]}(\mathbf{y}) = \mathbf{y}_{[z]}$, defining $\pi^\star_{[z]}(\mathbf{t}) = \sum_{j=1 j\neq z}^k \mathbf{1}_j t_j$ we get $\mathbf{t}_{[z]}^T \pi_{[z]}(\mathbf{y}) = \mathbf{y}^T \pi^\star_{[z]}(\mathbf{t})$. Then

$$\begin{aligned} \varphi_{\left(P^{(i,l)}_{x^{(1)},\dots,x^{(k)}} \circ \pi_{[z]}\right)}(\mathbf{t}_{[z]}) &= \int_{\mathbb{R}^k} e^{\mathrm{i}\mathbf{y}^T \pi^\star_{[z]}(\mathbf{t})} P^{(i,l)}_{x^{(1)},\dots,x^{(k)}}(\mathrm{d}\mathbf{y}) \\ &= \varphi_{P^{(i,l)}_{x^{(1)},\dots,x^{(k)}}}(\pi^\star_{[z]}(\mathbf{t})) \\ &= \varphi_{N_k(\mathbf{0},\Sigma(l,\mathbf{X}))}(\pi^\star_{[z]}(\mathbf{t})) \\ &= \exp\Big\{ -\frac{1}{2} \pi^\star_{[z]}(\mathbf{t})^T \Sigma(l, \mathbf{X}) \pi^\star_{[z]}(\mathbf{t}) \Big\} \\ &= \exp\Big\{ -\frac{1}{2} \mathbf{t}_{[z]}^T \widehat{\Sigma}(l, \mathbf{X}) \mathbf{t}_{[z]} \Big\}, \end{aligned}$$

where $\widehat{\Sigma}(l, \mathbf{X})$ is the matrix $\Sigma(l, \mathbf{X})$ without the z-th row and the z-th column. But since $\widehat{\Sigma}(l, \mathbf{X}) = \Sigma(l, \pi_{[z]}(x^{(1)}, \dots, x^{(k)}))$ we get $\varphi_{\left(P^{(i,l)}_{x^{(1)},\dots,x^{(k)}} \circ \pi_{[z]}\right)}(\mathbf{t}_{[z]}) = \varphi_{P^{(i,l)}_{\pi_{[z]}(x^{(1)},\dots,x^{(k)})}}(\mathbf{t}_{[z]})$ for each $\mathbf{t}_{[z]}$ and thus the two Fourier transformations coincides, as we wanted to prove.

SM E.3: Existence of a probability measure on $C(\mathbb{R}^I; \mathbb{R})$

If Z is, in addition, a continuous stochastic process then we will show that there exists a probability measure $\mathbb{P}^Z$ on $C(\mathbb{R}^I; \mathbb{R}) \subset \mathbb{R}^{\mathbb{R}^I}$ endowed with a σ-algebra $\mathcal{G} \subset \mathcal{B}(\mathbb{R}^{\mathbb{R}^I})$ such that the finite-dimensional distribution of Z' and Z coincide.

As suggested by Kallenberg (2002) (page 311) we consider $C(\mathbb{R}^I;\mathbb{R})$ with the topology of uniform convergence on compacts, that is

$$\begin{cases} \rho_{\mathbb{R}} : C(\mathbb{R}^I;\mathbb{R}) \times C(\mathbb{R}^I;\mathbb{R}) \to [0,\infty), \\ \qquad (\omega_1,\omega_2) \mapsto \rho_S(\omega_1,\omega_2) = \sum_{R=1}^{\infty} \frac{1}{2^R} \sup_{x\in\overline{B}_R(0)} \xi(|\omega_1(x)-\omega_2(x)|_{\mathbb{R}}) \end{cases} \tag{19}$$

The Borel σ-field $\mathcal{G} := \mathcal{B}(C(\mathbb{R}^I;\mathbb{R}),\rho_{\mathbb{R}})$ is generated by the evaluation maps π_x, thus it coincide with the product σ-field, i.e. $\mathcal{G} = \sigma(\Gamma)$, where

$$\Gamma = \left\{\Gamma_{x^{(1)},\dots,x^{(k)}}(A) | A = A_{x^{(1)}} \times \cdots \times A_{x^{(k)}}, A_{x^{(j)}} \in \mathcal{B}(\mathbb{R}), x^{(j)} \in \mathbb{R}^I, j \in \{1,\dots,k\}, k \in \mathbb{N}\right\}$$

where $\Gamma_{x^{(1)},\dots,x^{(k)}}(A) = \left\{\omega \in C(\mathbb{R}^I;\mathbb{R}) | \omega(x^{(1)}) \in A_{x^{(1)}}, \dots, \omega(x^{(k)}) \in A_{x^{(k)}}\right\}$. Note that since $\sigma(\Gamma) \subset \mathcal{B}(\mathbb{R}^{\mathbb{R}^I})$ then $\mathcal{G} = \sigma(\Gamma) \subset \mathcal{B}(\mathbb{R}^{\mathbb{R}^I})$.

Theorem 3. *There exists a unique probability measure $\mathbb{P}^Z$ on $(C(\mathbb{R}^I;\mathbb{R}),\mathcal{G})$ such that the canonical process Z' restricted to $(C(\mathbb{R}^I;\mathbb{R}),\mathcal{G}))$ has finite-dimensional distributions that coincide with those of Z.*

For the existence of $\mathbb{P}^Z$ consider the following

Lemma 6. *Let $(Z_x)_{x\in\mathbb{R}^I}$ be a $\mathbb{R}$-valued continuous stochastic process defined on $(\Omega,\mathcal{H},\mathbb{P})$. Then*

$$\begin{cases} \mathcal{Z} : \Omega \to C(\mathbb{R}^I;\mathbb{R}) \\ \qquad \omega \to \mathcal{Z}(\omega) = (Z_x(\omega))_{x\in\mathbb{R}^I} \end{cases}$$

is a random variable, i.e. measurable from $(\Omega,\mathcal{H})$ into $(C(\mathbb{R}^I;\mathbb{R}),\mathcal{G})$.

Proof. By previous proposition $\mathcal{G} = \sigma(\Gamma)$, then taking $\mathcal{O} \in \sigma(\Gamma)$, $\mathcal{O} = \Gamma_{x^{(1)},\dots,x^{(k)}}(A)$ for some $k \in \mathbb{N}$, $\{x^{(1)},\dots,x^{(k)}\} \subset \mathbb{R}^I$ and $A = A_{x^{(1)}} \times \cdots \times A_{x^{(k)}}, A_{x^{(j)}} \in \mathcal{B}(\mathbb{R})$,we get

$$\begin{aligned} \{\omega \in \Omega | \mathcal{Z}(\omega) \in \mathcal{O}\} &= \{\omega \in \Omega | Z_{x^{(1)}}(\omega) \in A_{x^{(1)}}, \dots, Z_{x^{(k)}}(\omega) \in A_{x^{(k)}}\} \\ &= \bigcap_{j=1}^{k} \{Z_{x^{(j)}} \in A_{x^{(j)}}\} \in \mathcal{H} \end{aligned}$$

where we used that $Z_{x^{(j)}}$ are random variables from $(\Omega,\mathcal{H})$ into $(\mathbb{R},\mathcal{B}(\mathbb{R}))$. $\square$

Then we can define a probability measure $\mathbb{P}^Z$ on $(C(\mathbb{R}^I;\mathbb{R}),\mathcal{G})$ being the image measure of $\mathcal{Z}$ under $\mathbb{P}$, that is

$$\forall \mathcal{O} \in \mathcal{G}, \quad \mathbb{P}^Z(\mathcal{O}) = \mathbb{P}(\mathcal{Z} \in \mathcal{O})$$

Now we prove that the finite-dimensional distributions of Z' coincide with those of Z. It is sufficient to prove the following

Lemma 7. *$\mathbb{P}^Z$ coincide wit the image measure of the canonical process Z' under $\mathbb{P}'$ restricted to $(C(\mathbb{R}^I;\mathbb{R}),\mathcal{G})$.*

Proof. Fix $\mathcal{O} \in \mathcal{G} = \sigma(\Gamma)$, $\mathcal{O} = \Gamma_{x^{(1)},\dots,x^{(k)}}(A)$ for some $k \in \mathbb{N}$, $\{x^{(1)},\dots,x^{(k)}\} \subset \mathbb{R}^I$ and $A = A_{x^{(1)}} \times \cdots \times A_{x^{(k)}}, A_{x^{(j)}} \in \mathcal{B}(\mathbb{R})$. By definition of $\mathbb{P}^Z$,

$$\begin{aligned} \mathbb{P}^Z(\mathcal{O}) &= \mathbb{P}(\mathcal{Z} \in \Gamma_{x^{(1)},\dots,x^{(k)}}(A)) \\ &= \mathbb{P}(\{\omega \in \Omega | \mathcal{Z}(\omega) \in \mathcal{O}\}) \\ &= \mathbb{P}(\{\omega \in \Omega | Z_{x^{(1)}}(\omega) \in A_{x^{(1)}}, \dots, Z_{x^{(k)}}(\omega) \in A_{x^{(k)}}) \\ &= P^Z_{x^{(1)},\dots,x^{(k)}}(A) \end{aligned}$$

By Daniell-Kolmogorv extension theorem the finite-dimensional distributions of Z coincide with those of the canonical process Z' under $\mathbb{P}'$, then $P^Z_{x^{(1)},\dots,x^{(k)}}(A) = \mathbb{P}'(Z' \in \mathcal{O})$. $\square$

The uniqueness of $\mathbb{P}^Z$ follows by the uniqueness $\mathbb{P}'$.

SM E.4: $\sigma(\times_{i=1}^{\infty} C(\mathbb{R}^I;\mathbb{R})) \subset \sigma(C(\mathbb{R}^I;\mathbb{R}^\infty)$

First, note that $\times_{i=1}^{\infty} C(\mathbb{R}^I;\mathbb{R}) \simeq C(\mathbb{R}^I;\mathbb{R}^\infty)$, indeed the map

$$\Xi : C(\mathbb{R}^I;\mathbb{R}^\infty) \to \times_{i=1}^{\infty} C(\mathbb{R}^I;\mathbb{R}), \qquad \omega \mapsto (\omega_1, \omega_2, \dots)$$

is an isomorphism because is linear and bijective, indeed ω is $\|\cdot\|_\infty$-continuous if and only if each component ω_i is $|\cdot|$-continuous. It means that each element in one space could be seen as an element in the other and vice-versa, but different topologies are defined on these spaces. Now we prove that the sigma algebra generated by the product topology in $\times_{i=1}^{\infty} C(\mathbb{R}^I;\mathbb{R})$ is contained on the sigma algebra generated by the topology of uniform convergence on compact set in $C(\mathbb{R}^I;\mathbb{R}^\infty)$. For each $f, g \in C(\mathbb{R}^I;\mathbb{R}^\infty)$ we have the following distances

$$\begin{cases} \rho_{prod}(f,g) = \sum_{i=1}^{\infty} \frac{1}{2^i} \xi\Big(\sum_{R=1}^{\infty} \frac{1}{2^R} \sup_{x \in B_R(0)} \xi(|f_i(x) - g_i(x)|)\Big), & \text{on } \times_{i=1}^{\infty} C(\mathbb{R}^I;\mathbb{R}) \\ \rho_{unif}(f,g) = \sum_{R=1}^{\infty} \frac{1}{2^R} \sup_{x \in B_R(0)} \xi\Big(\sum_{i=1}^{\infty} \frac{1}{2^i} \xi(|f_i(x) - g_i(x)|)\Big), & \text{on } C(\mathbb{R}^I, \mathbb{R}^\infty) \end{cases} \tag{20}$$

Using that ξ is increasing and continuous and that $\sup_x(\sum_i h_i(x)) \le \sum_i \sup_x h_i(x)$ it can be proved that there exists a constant $C > 0$ such that $\|f\|_{unif} \le C\|f\|_{prod}$. This mean that if $h \in B_\epsilon^{prod}(f) = \{g : \|f - g\|_{prod} < \epsilon\}$ than $h \in B_{C\epsilon}^{unif}(f) = \{g : \|f - g\|_{unif} < \epsilon\}$, that is $B_\epsilon^{prod}(f) \subset B_{C\epsilon}^{unif}(f)$ which implies $\sigma(\rho_{prod}) \subset \sigma(\rho_{unif})$. In particular each compact with respect to $\|\cdot\|_{prod}$ is compact with respect to $\|\cdot\|_{unif}$, indeed considering a $\|\cdot\|_{prod}$-compact K then for every sequence $(k_i) \subset K$ there exists $(k_{i_j}) \subset K$ and $k \in K$ such that $\|k_{i_j} - k\|_{prod} \to 0$. Moreover $\|k_{i_j} - k\|_{unif} \le C\|k_{i_j} - k\|_{prod} \to 0$, i.e. K is compact with respect to $\|\cdot\|_{unif}$.

SM F

In this section we prove the Proposition 1.

Proof. By Proposition 16.6 of Kallenberg (2002) $f(n) \xrightarrow{d} f$ in $C(\mathbb{R}^I;S)$ iff $f(n) \xrightarrow{d} f$ in $C(K;S)$ for any $K \subset \mathbb{R}^I$ compact. By Lemma 16.2 of of Kallenberg (2002) the latter holds iff $f(n) \xrightarrow{f_d} f$ and $(f(n))_{n\ge1}$ is relatively compact in distribution in $C(K;S)$. Note that converge of the finite-dimensional distributions holds in $\mathbb{R}^I$ iif it holds in the restriction K for any compact $K \subset \mathbb{R}^I$. The space $(C(K;S), \rho_K)$, namely the space of continuous functions from a generic compact $K \subset \mathbb{R}^I$ to a Polish space S and $C(K;S)$ endowed with the uniform metric $\rho_K(f,g) = \sup_{x\in K} d(f(x), g(x))$, is itself a Polish space (Aliprantis & Border, 2006, Lemma 3.97 and Lemma 3.99). Thus by Proposition 16.3 of Kallenberg (2002), i.e. Prohorov Theorem, on $C(K,S)$ $(f(n))_{n\ge1}$ is relatively compact in distribution iif $(f(n))_{n\ge1}$ is uniformly tight. Thus, so far we have shown that $f(n) \xrightarrow{d} f$ in $C(\mathbb{R}^I;S)$ iff: i) $f(n) \xrightarrow{f_d} f$ and ii) the sequence $(f(n))_{n\ge1}$ is uniformly tight on $C(K,S)$ for a generic compact $K \subset \mathbb{R}^I$. It remains to show that the latter holds if $(f(n))_{n\ge1}$ is uniformly tight on $C(\mathbb{R}^I;S)$. Fix K compact in $\mathbb{R}^I$ and Consider the map

$$\pi_K : (C(\mathbb{R}^I;S), \rho_S) \to (C(K;S), \rho_K), \quad f \to f_{|K}$$

where $f_{|K}$ is the restriction of f to K and ρ_S is the metric $\rho_\mathbb{R}$ defined in (19) when $S = \mathbb{R}$ and ρ_{unif} defined in (20) when $S = \mathbb{R}^\infty$. By proposition 16.4 of Kallenberg (2002) if π_K is continuous then it moves uniformly tight sequences into uniformly tight sequences. The continuity of π_K follows by the proof of Proposition 16.6 of Kallenberg (2002). $\square$