A.

Proposition 1. *Suppose that* $\{k^l(\boldsymbol{x},\boldsymbol{x}') = \psi^l(\boldsymbol{x}-\boldsymbol{x}')\}_{l=1}^L$ *on* $\mathbb{R}^d$ *are bounded, continuous reproducing kernels. Let* $P^l \triangleq \mathbb{P}(\boldsymbol{X}^l|\boldsymbol{X}^{1:l-1})$ *for* $l=1,\dots,L$ *with* $P^1=\mathbb{P}(\boldsymbol{X}^1)$*. Then for any* $\mathbb{P}(\boldsymbol{X}^1,\dots,\boldsymbol{X}^L), \mathbb{Q}(\boldsymbol{X}'^1,\dots,\boldsymbol{X}'^L) \in M_+^1(\times_{l=1}^L \mathcal{X}^l)$,

$$\mathcal{S}_{JRS}(\mathbb{P},\mathbb{Q}) = \prod_{l=1}^L \langle \phi_{P^l}(\omega), \phi_{Q^l}(\omega)\rangle_{L^2(\mathbb{R}^d,\Lambda^l)}, \tag{1}$$

where $\phi_{P^l}(\omega)$ *and* $\phi_{Q^l}(\omega)$ *are the characteristic functions of the distributions* P^l *and* Q^l*, and* Λ^l *is a (normalized) non-negative Borel measure characterized by* $\psi^l(\boldsymbol{x}-\boldsymbol{x}')$.

Proof. The proof of Proposition 1 is a simple corollary of the famous Bochner's theorem (Bochner, 1959).

Lemma 1 (Bochner's theorem (Bochner, 1959; Wendland, 2004))**.** *A complex-valued bounded continuous kernel* $k(\boldsymbol{x},\boldsymbol{x}') = \psi(\boldsymbol{x}-\boldsymbol{x}')$ *on* $\mathbb{R}^d$ *is positive definite if and only if it is the Fourier transform of a finite non-negative Borel measure* Λ *on* $\mathbb{R}^d$*, i.e.,*

$$\psi(\mathbf{x}-\mathbf{x}') = \int_{\mathbb{R}^d} e^{-i\omega^\top(\mathbf{x}-\mathbf{x}')}\mathrm{d}\Lambda(\omega)$$

By the definition of joint representation similarity in Definition 4, we have,

$$\begin{aligned}
\mathcal{S}_{JRS}(\mathbb{P},\mathbb{Q}) &= \langle \mathcal{C}_{\mathbf{X}^{1:L}}(\mathbb{P}), \mathcal{C}_{\mathbf{X}'^{1:L}}(\mathbb{Q})\rangle_{\otimes_{l=1}^L \mathcal{RKHS}^l} \\
&= \langle \int_{\times_{l=1}^L \mathbf{X}^l} (\otimes_{l=1}^L k^l(\cdot,\mathbf{x}^l))\mathrm{d}\mathbb{P}(\mathbf{x}^1,\dots,\mathbf{x}^L), \int_{\times_{l=1}^L \mathbf{X}'^l} (\otimes_{l=1}^L k^l(\cdot,\mathbf{x}'^l))\mathrm{d}\mathbb{Q}(\mathbf{x}'^1,\dots,\mathbf{x}'^L)\rangle_{\otimes_{l=1}^L \mathcal{RKHS}^l} \\
&= \int_{\times_{l=1}^L \mathbf{X}^l}\int_{\times_{l=1}^L \mathbf{X}'^l} \langle \otimes_{l=1}^L k^l(\cdot,\mathbf{x}^l), \otimes_{l=1}^L k^l(\cdot,\mathbf{x}'^l)\rangle_{\otimes_{l=1}^L \mathcal{RKHS}^l}\mathrm{d}\mathbb{P}(\mathbf{x}^1,\dots,\mathbf{x}^L)\mathrm{d}\mathbb{Q}(\mathbf{x}'^1,\dots,\mathbf{x}'^L) \\
&= \int_{\times_{l=1}^L \mathbf{X}^l}\int_{\times_{l=1}^L \mathbf{X}'^l} \prod_{l=1}^L \langle k^l(\cdot,\mathbf{x}^l), k^l(\cdot,\mathbf{x}'^l)\rangle_{\mathcal{RKHS}^l}\mathrm{d}\mathbb{P}(\mathbf{x}^1,\dots,\mathbf{x}^L)\mathrm{d}\mathbb{Q}(\mathbf{x}'^1,\dots,\mathbf{x}'^L) \\
&\overset{(a)}{=} \int_{\times_{l=1}^L \mathbf{X}^l}\int_{\times_{l=1}^L \mathbf{X}'^l} \prod_{l=1}^L k^l(\mathbf{x}^l,\mathbf{x}'^l)\mathrm{d}\mathbb{P}(\mathbf{x}^1,\dots,\mathbf{x}^L)\mathrm{d}\mathbb{Q}(\mathbf{x}'^1,\dots,\mathbf{x}'^L) \\
&\overset{(b)}{=} \int_{\times_{l=1}^L \mathbf{X}^l}\int_{\times_{l=1}^L \mathbf{X}'^l} \prod_{l=1}^L \int_{\mathbb{R}^d} e^{-i\omega^\top(\mathbf{x}^l-\mathbf{x}'^l)}\mathrm{d}\Lambda^l(\omega)\mathrm{d}\mathbb{P}(\mathbf{x}^1,\dots,\mathbf{x}^L)\mathrm{d}\mathbb{Q}(\mathbf{x}'^1,\dots,\mathbf{x}'^L) \\
&= \int_{\times_{l=2}^L \mathbf{X}^l}\int_{\mathbf{X}^1}\mathrm{d}\mathbb{P}(\mathbf{x}^1)\int_{\times_{l=2}^L \mathbf{X}'^l}\int_{\mathbf{X}'^1}\mathrm{d}\mathbb{Q}(\mathbf{x}'^1)\int_{\mathbb{R}^d} e^{-i\omega^\top(\mathbf{x}^1-\mathbf{x}'^1)}\mathrm{d}\Lambda^1(\omega)\prod_{l=2}^L\int_{\mathbb{R}^d} e^{-i\omega^\top(\mathbf{x}^l-\mathbf{x}'^l)}\mathrm{d}\Lambda^l(\omega)\mathrm{d}\mathbb{P}(\mathbf{x}^{2:L}|\mathbf{x}^1)\mathrm{d}\mathbb{Q}(\mathbf{x}'^{2:L}|\mathbf{x}'^1) \\
&\overset{(c)}{=} \int_{\times_{l=2}^L \mathbf{X}^l}\int_{\times_{l=2}^L \mathbf{X}'^l}\int_{\mathbb{R}^d}\int e^{-i\omega^\top\mathbf{x}^1}\mathrm{d}\mathbb{P}(\mathbf{x}^1)\int e^{i\omega^\top\mathbf{x}'^1}\mathrm{d}\mathbb{Q}(\mathbf{x}'^1)\mathrm{d}\Lambda^1(\omega)\prod_{l=2}^L\int_{\mathbb{R}^d} e^{-i\omega^\top(\mathbf{x}^l-\mathbf{x}'^l)}\mathrm{d}\Lambda^l(\omega)\mathrm{d}\mathbb{P}(\mathbf{x}^{2:L}|\mathbf{x}^1)\mathrm{d}\mathbb{Q}(\mathbf{x}'^{2:L}|\mathbf{x}'^1) \\
&= \int_{\times_{l=2}^L \mathbf{X}^l}\int_{\times_{l=2}^L \mathbf{X}'^l}\int_{\mathbb{R}^d}\overline{\phi_{P^1}(\omega)}\phi_{Q^1}(\omega)\mathrm{d}\Lambda^1(\omega)\prod_{l=2}^L\int_{\mathbb{R}^d} e^{-i\omega^\top(\mathbf{x}^l-\mathbf{x}'^l)}\mathrm{d}\Lambda^l(\omega)\mathrm{d}\mathbb{P}(\mathbf{x}^{2:L}|\mathbf{x}^1)\mathrm{d}\mathbb{Q}(\mathbf{x}'^{2:L}|\mathbf{x}'^1) \\
&= \int_{\times_{l=2}^L \mathbf{X}^l}\int_{\times_{l=2}^L \mathbf{X}'^l}\langle\phi_{P^1}(\omega),\phi_{Q^1}(\omega)\rangle_{L^2(\mathbb{R}^d,\Lambda^1)}\prod_{l=2}^L\int_{\mathbb{R}^d} e^{-i\omega^\top(\mathbf{x}^l-\mathbf{x}'^l)}\mathrm{d}\Lambda^l(\omega)\mathrm{d}\mathbb{P}(\mathbf{x}^{2:L}|\mathbf{x}^1)\mathrm{d}\mathbb{Q}(\mathbf{x}'^{2:L}|\mathbf{x}'^1) \\
&= \langle\phi_{P^1}(\omega),\phi_{Q^1}(\omega)\rangle_{L^2(\mathbb{R}^d,\Lambda^1)}\int_{\times_{l=2}^L \mathbf{X}^l}\int_{\times_{l=2}^L \mathbf{X}'^l}\prod_{l=2}^L\int_{\mathbb{R}^d} e^{-i\omega^\top(\mathbf{x}^l-\mathbf{x}'^l)}\mathrm{d}\Lambda^l(\omega)\mathrm{d}\mathbb{P}(\mathbf{x}^{2:L}|\mathbf{x}^1)\mathrm{d}\mathbb{Q}(\mathbf{x}'^{2:L}|\mathbf{x}'^1), \\
&\overset{(d)}{=} \prod_{l=1}^L\langle\phi_{P^l}(\omega),\phi_{Q^l}(\omega)\rangle_{L^2(\mathbb{R}^d,\Lambda^l)},
\end{aligned}$$

where (a) is obtained by invoking the reproducing property, (b) is obtained by invoking Bochner's theorem, (c) is obtainied by invoking Fubini's theorem, (d) is obtained by the principle of mathematical reduction. $\square$