The Wasserstein metric, Wasserstein-Fréchet mean, simulation results and additional proofs

Alexander Petersen and Hans-Georg Müller

S.1. The Wasserstein Metric. The equivalence of the metrics

\[ d_Q(f,g)^2 = \int_0^1 (F^{-1}(t) - G^{-1}(t))^2 \, dt \quad \text{and} \quad d_W(f,g)^2 = \inf_{X \sim f, Y \sim g} E(X - Y)^2 \]

is well known. It can be easily seen by applying a covariance identity due to [28]. If \( X \sim F, Y \sim G \) and \( (X,Y) \sim H \), then this identity states that

\[ \text{Cov}(X,Y) = \iint \{H(u,v) - F(u)G(v)\} \, du \, dv. \]

Expanding the expectation \( E(X - Y)^2 \), one finds that the distance is obtained by maximizing \( E(XY) \), or, equivalently, by maximizing \( \text{Cov}(X,Y) \). For a random variable \( U \) that is uniformly distributed on \([0,1]\), take \( X^* = F^{-1}(U) \) and \( Y^* = G^{-1}(U) \). Then \( X^* \sim F, Y^* \sim G \) and the distribution function of \((X^*,Y^*)\) is given by \( H^*(u,v) = \min(F(u),G(v)) \). Clearly, for any joint distribution of \( X \sim F \) and \( Y \sim G \), we have \( H \leq H^* \). By Hoeffding’s inequality, this means \( \text{Cov}(X,Y) \leq \text{Cov}(X^*,Y^*) \). Thus,

\[ d_W(f,g)^2 = E[(X^* - Y^*)^2] = E[(F^{-1}(U) - G^{-1}(U))^2] \]

\[ = \int_0^1 (F^{-1}(t) - G^{-1}(t))^2 \, dt. \]
Let $Q$ be the quantile process corresponding to the density process $f \sim \tilde{F}$ and set $Q_\oplus(t) = E(Q(t))$. For $q_\oplus = Q'_\oplus$ and $F_\oplus = Q^{-1}_\oplus$, the Wasserstein-Fréchet mean is

$$f_\oplus(x) = \frac{1}{q_\oplus(F_\oplus(x))}.$$ 

Its estimation can thus be reduced to estimating the function $q_\oplus$. Due to the restrictions on the space $\mathcal{F}$ (see assumption (A1)), we can pass differentiation inside the expectation so that $E(Q'(t)) = q_\oplus(t)$. This suggests averaging the quantile densities of the sample to obtain an estimator for $q_\oplus$.

Starting with either the densities, $f_i$, or their estimates, $\tilde{f}_i$, $i = 1, \ldots, n$, we therefore use the corresponding quantile densities ($q_i$ or $\tilde{q}_i$) to estimate $q_\oplus$ by

$$\tilde{q}_\oplus(t) = \frac{1}{n} \sum_{i=1}^{n} q_i(t), \quad \text{respectively,} \quad \hat{q}_\oplus(t) = \frac{1}{n} \sum_{i=1}^{n} \tilde{q}_i(t).$$

Computing the corresponding distribution functions, we thus estimate the Wasserstein-Fréchet mean by

$$\tilde{f}_\oplus(x) = \frac{1}{\tilde{q}_\oplus(\tilde{F}_\oplus(x))}, \quad \text{respectively,} \quad \hat{f}_\oplus(x) = \frac{1}{\hat{q}_\oplus(\hat{F}_\oplus(x))}.$$

As Theorem 2 requires a rate of convergence $\gamma_n$ for the Wasserstein-Fréchet mean estimator based on fully observed densities, the following result shows that we make take $\gamma_n = n^{-1/2}$ in the case of fully observed densities.

**Proposition 3.** Under assumption (A1), the estimator $\tilde{f}_\oplus$ of $f_\oplus$ for the Wasserstein-Fréchet mean satisfies

$$d_W(f_\oplus, \tilde{f}_\oplus) = O_p(n^{-1/2}).$$

**Proof.** By Thm 3.9 in [9], $d_2(q_\oplus, \tilde{q}_\oplus) = O_p(n^{-1/2})$. As $|Q_\oplus(t) - \tilde{Q}_\oplus(t)| \leq d_2(q_\oplus, \tilde{q}_\oplus)$, we also have

$$d_W(f_\oplus, \tilde{f}_\oplus) = d_2(Q_\oplus, \tilde{Q}_\oplus) = O_p(n^{-1/2}).$$

$\square$
S.2. Simulation Results for the Wasserstein Metric. Figure 7 shows the distribution of fraction of variance explained values in terms of the distance $d_W$ for all simulation settings, similar to Figure 2 in the main text which shows the results for the ordinary $L^2$ distance. The use of the Wasserstein distance more clearly demonstrates the weakness of ordinary FPCA. The Hilbert sphere method performs relatively better in the context of metric $d_W$ than the $L^2$ metric, but is still outperformed by the transformation method using the log quantile density transformation, $\psi_Q$.

Fig 7: Boxplots of fraction of variance explained for 200 simulations, using the Wasserstein metric, $d_W$. The first row corresponds to fully observed densities and the second corresponds to estimated densities. The columns correspond to settings 1, 2 and 3 from left to right (see Table 1). The methods are denoted by ‘FPCA’ for ordinary FPCA on the densities, ‘LQD’ for the transformation approach with $\psi_Q$ and ‘HS’ for the Hilbert sphere method.
S.3. Listing of All Assumptions. The following is a systematic compilation of all assumptions, subsets of which are used for various results and some of which have been stated in the main text. Recall that $d_2$ and $d_\infty$ denote the $L^2$ and uniform metrics, respectively, and $\|\cdot\|_2$ and $\|\cdot\|_\infty$ denote the corresponding norms.

(A1) For all $f \in \mathcal{F}$, $f$ is continuously differentiable. Moreover, there is a constant $M > 1$ such that, for all $f \in \mathcal{F}$, $\|f\|_\infty$, $\|1/f\|_\infty$ and $\|f'\|_\infty$ are all bounded above by $M$.

(D1) For a sequence $b_N = o(1)$, the density estimator $\hat{f}$, based on an i.i.d. sample of size $N$, satisfies $\hat{f} \geq 0$, $\int_{0}^{1} \hat{f}(x) \, dx = 1$ and

$$\sup_{f \in \mathcal{F}} E(d_2(f, \hat{f})^2) = O(b_N^2).$$

(D2) For a sequence $a_N = o(1)$ and some $R > 0$, the density estimator $\hat{f}$, based on an i.i.d. sample of size $N$, satisfies

$$\sup_{f \in \mathcal{F}} P(d_\infty(f, \hat{f}) > Ra_N) \to 0.$$

(S1) Let $\hat{f}$ be a density estimator that satisfies (D2), and suppose densities $f_i \in \mathcal{F}$ are estimated by $\hat{f}_i$ from i.i.d. samples of size $N_i = N_i(n)$, $i = 1, \ldots, n$, respectively. There exists a sequence of lower bounds $m(n) \leq \min_{1 \leq i \leq n} N_i$ such that $m(n) \to \infty$ as $n \to \infty$ and

$$n \sup_{f \in \mathcal{F}} P(d_\infty(f, \hat{f}) > Ra_m) \to 0$$

where, for generic $f \in \mathcal{F}$, $\hat{f}$ is the estimated density from a sample of size $N(n) \geq m(n)$.

(K1) The kernel $\kappa$ is of bounded variation and is symmetric about 0.

(K2) The kernel $\kappa$ satisfies $\int_{0}^{1} \kappa(u) \, du > 0$, and $\int_{\mathbb{R}} |u|\kappa(u) \, du$, $\int_{\mathbb{R}} \kappa^2(u) \, du$ and $\int_{\mathbb{R}} |u|\kappa^2(u) \, du$ are finite.

(T0) Let $f, g \in \mathcal{G}$ with $f$ differentiable and $\|f'\|_\infty < \infty$. Set

$$D_0 \geq \max \left( \|f\|_\infty, \|1/f\|_\infty, \|g\|_\infty, \|1/g\|_\infty, \|f'\|_\infty \right).$$

There exists $C_0$ depending only on $D_0$ such that

$$d_2(\psi(f), \psi(g)) \leq C_0 d_2(f, g), \quad d_\infty(\psi(f), \psi(g)) \leq C_0 d_\infty(f, g).$$

(T1) Let $f \in \mathcal{G}$ be differentiable with $\|f'\|_\infty < \infty$ and let $D_1$ be a constant bounded below by $\max (\|f\|_\infty, \|1/f\|_\infty, \|f'\|_\infty)$. Then $\psi(f)$ is differentiable and there exists $C_1 > 0$ depending only on $D_1$ such that $\|\psi(f)\|_\infty \leq C_1$ and $\|\psi(f)'\|_\infty \leq C_1$. 
(T2) Let \( d \) be the selected metric in density space, \( Y \) be continuous and \( X \) be differentiable on \( T \) with \( \| X' \|_\infty < \infty \). There exist constants \( C_2 = C_2(\| X \|_\infty, \| X' \|_\infty) > 0 \) and \( C_3 = C_3(d_\infty(X,Y)) > 0 \) such that
\[
d(\psi^{-1}(X), \psi^{-1}(Y)) \leq C_2 C_3 d_2(X,Y)
\]
and, as functions, \( C_2 \) and \( C_3 \) are increasing in their respective arguments.

(T3) For a given metric \( d \) on the space of densities and \( f_1,K = f_1(\cdot,K,\psi) \) (see (4.5)), \( V_\infty - V_K \to 0 \) and \( E(d(f,f_1,K)^4) = O(1) \) as \( K \to \infty \).

S.4. Additional Proofs.

**Lemma 1.** Let \( A \) be a closed and bounded interval of length \( |A| \) and assume \( X : A \to \mathbb{R} \) is continuous with Lipschitz constant \( L \). Then
\[
\| X \|_\infty \leq 2 \max \left( |A|^{-1/2} \| X \|_2, \ L^{1/3} \| X \|_2^{2/3} \right).
\]

**Proof of Lemma 1.** Let \( t^* \) satisfy \( |X(t^*)| = \| X \|_\infty \) and define \( I = [t^* - \| X \|_\infty/(2L), t^* + \| X \|_\infty/(2L)] \cap A \). Then, for \( t \in I \), \( |X(t)| \geq \| X \|_\infty/2 \).
If \( I = A \),
\[
\| X \|_2^2 = \int_A X^2(s) \, ds \geq \frac{|A|}{4} \| X \|_\infty^2,
\]
so \( \| X \|_\infty \leq 2|A|^{-1/2} \| X \|_2 \). If \( I \neq A \), suppose without loss of generality that \( t^* + \| X \|_\infty/(2L) \in A \). Then
\[
\| X \|_2^2 \geq \int_{t^*}^{t^* + \| X \|_\infty/(2L)} X^2(s) \, ds \geq \frac{\| X \|_\infty^2}{4} \cdot \frac{\| X \|_\infty}{2L} = \frac{\| X \|_\infty^3}{8L},
\]
so \( \| X \|_\infty \leq 2L^{1/3} \| X \|_2^{2/3} \). \( \square \)

**Lemma 2.** Let \( X \) be a stochastic process on a closed interval \( T \subset \mathbb{R} \) such that \( \| X \|_\infty < C \) and \( \| X' \|_\infty < C \) almost surely. Let \( \nu \) and \( H \) be the mean and covariance functions associated with \( X \), and \( \rho_k \) and \( \tau_k \), \( k \geq 1 \), be the eigenfunctions and eigenvalues of the integral operator with kernel \( H \). Then
\[
\| \nu \|_\infty < C, \ \| H \|_\infty < 4C^2 \text{ and } \| \rho_k \|_\infty < 4C^2 |T|^{1/2} \tau_k^{-1} \text{ for all } k \geq 1.
\]
Additionally, \( \| \nu' \|_\infty < C \) and \( \| \rho'_k \|_\infty < 4C^2 |T|^{1/2} \tau_k^{-1} \) for all \( k \geq 1 \).

**Proof.** Since the bounds on \( X \) and \( X' \) are deterministic, \( \| \nu \|_\infty \) and \( \| H \|_\infty \) are both bounded by the given constants. The bound on \( \| \rho_k \|_\infty \) follows since \( \rho_k(t) = \tau_k^{-1} \int_T H(s,t) \rho_k(s) \, ds \) and \( \| \rho_k \|_2 = 1 \). Dominated convergence implies that \( \nu' \) exists and is bounded by \( C \), and also implies the bound.
of $4C^2$ for the partial derivatives of $H$, which then leads to the bounds on $\rho'_k$ for all $k$.

\begin{lemma}
Under assumptions (A1) and (T1), with $\hat{\nu}, \tilde{\nu}, \hat{H}, \tilde{H}$ as in (4.2) and (4.3),
\begin{equation}
\begin{aligned}
d_2(\nu, \hat{\nu}) &= O_p(n^{-1/2}), &\quad d_2(H, \hat{H}) &= O_p(n^{-1/2}), \\
d_\infty(\nu, \hat{\nu}) &= O_p\left(\left(\frac{\log n}{n}\right)^{1/2}\right), &\quad d_\infty(H, \hat{H}) &= O_p\left(\left(\frac{\log n}{n}\right)^{1/2}\right).
\end{aligned}
\end{equation}
\end{lemma}

Under the additional assumptions (D1), (D2) and (S1), we have
\begin{equation}
\begin{aligned}
d_2(\nu, \hat{\nu}) &= O_p(n^{-1/2} + b_m), &\quad d_2(H, \hat{H}) &= O_p(n^{-1/2} + b_m), \\
d_\infty(\nu, \hat{\nu}) &= O_p\left(\left(\frac{\log n}{n}\right)^{1/2} + a_m\right), &\quad d_\infty(H, \hat{H}) &= O_p\left(\left(\frac{\log n}{n}\right)^{1/2} + a_m\right).
\end{aligned}
\end{equation}

\begin{proof}
Assumptions (A1) and (T1) imply $E\|X\|_2^2 < \infty$, so the first line in (S.1) follows from Theorems 3.9 and 4.2 in [9]. The second line in (S.1) follows from Corollaries 2.3(b) and 3.5(b) in [33]. We will show the argument for the mean estimate in (S.2), and the covariance follows similarly.

Let $M$ be as given in assumption (A1) and set $D_1 = 2M$. Define
\[ E_n = \bigcap_{i=1}^{n} \{ d_\infty(f_i, \hat{f}_i) \leq D_1^{-1} \} . \]
Then $P(E_n) \to 0$ by assumptions (D2) and (S1). Take $C_1$ as given in (T1) for $D_1$ as defined above. Also by (S1), there is $R > 0$ such that
\[ P(\{ d_\infty(\hat{\nu}, \nu) > Ra_m \} \cap E_n) \leq n \max_{1 \leq i \leq n} P(d_\infty(f_i, \hat{f}_i) > C_1^{-1}Ra_m) \to 0 \]
as $n \to \infty$, so $d_\infty(\hat{\nu}, \nu) = O_p(a_m)$. Thus, by the triangle inequality, $d_\infty(\nu, \hat{\nu}) = O_p\left(\left(\frac{\log n}{n}\right)^{1/2} + a_m\right)$.
\end{proof}
Next, letting \( \hat{X}_i = \psi(\hat{f}_i) \),

\[
P \left( \{d_2(\hat{\nu}, \hat{\nu}) > R \} \cap E_n \right) \leq P \left( \sum_{i=1}^{n} d_2(X_i, \hat{X}_i) > Rn \right) \cap E_n \\
\leq P \left( \sum_{i=1}^{n} d_2(f_i, \hat{f}_i) > C_1^{-1}Rn \right) \\
\leq C_1 R^{-1} n^{-1} \sum_{i=1}^{n} \sqrt{\mathbb{E} (d_2(f_i, \hat{f}_i)^2)} = R^{-1} O(b_m),
\]

which shows that \( d_2(\hat{\nu}, \hat{\nu}) = O_p(b_m) \), so the result holds by the triangle inequality. \( \square \)

**Corollary 1.**  Under assumption (A1) and (T1), letting \( A_k = \|\rho_k\|_\infty \), with \( \delta_k \) as in (5.1),

\[
|\tau_k - \hat{\tau}_k| = O_p(n^{-1/2}), \\
\quad d_2(\rho_k, \hat{\rho}_k) = \delta_k^{-1} O_p(n^{-1/2}), \quad \text{and} \\
\quad d_\infty(\rho_k, \hat{\rho}_k) = \delta_k^{-1} O_p \left( \frac{(\log n)^{1/2} + \delta_k^{-1} + A_k}{n^{1/2}} \right),
\]

(S.3)

where all \( O_p \) terms are uniform over \( k \). If the additional assumptions (D1), (D2) and (S1) hold,

\[
|\tau_k - \hat{\tau}_k| = O_p(n^{-1/2} + b_m), \\
\quad d_2(\rho_k, \hat{\rho}_k) = \delta_k^{-1} O_p(n^{-1/2} + b_m), \quad \text{and} \\
\quad d_\infty(\rho_k, \hat{\rho}_k) = \hat{\tau}_k^{-1} O_p \left( \frac{(\log n)^{1/2} + \delta_k^{-1} + A_k}{n^{1/2}} + a_m + b_m[\delta_k^{-1} + A_k] \right),
\]

(S.4)

where again all \( O_p \) terms are uniform over \( k \).

**Proof.** First, observe that (A1) and (T1) together imply that \( X \) satisfies the assumptions of Lemma 2. The first two lines of both (S.3) and (S.4) follow by applying Lemmas 4.2 and 4.3 of [9] with the rates given in Lemma 3, above. For the uniform metric on the eigenfunctions, we follow the argument given in the proof of Lemma 1 in [36] to find that

\[
d_\infty(\tau_k \rho_k, \hat{\tau}_k \hat{\rho}_k) \leq |T|^{1/2} \left[ d_\infty(H, \hat{H}) + \|H\|_\infty d_2(\rho_k, \hat{\rho}_k) \right] = O_p \left( \frac{(\log n)^{1/2} + \delta_k^{-1}}{n^{1/2}} \right),
\]
It follows that
\[ |\rho_k(s) - \tilde{\rho}_k(s)| \leq \tilde{\tau}_k^{-1} (|\tau_k \rho_k(s) - \tilde{\tau}_k \rho_k(s)| + |\rho_k(s)| |\tau_k - \tilde{\tau}_k|) \]
\[ = \tilde{\tau}_k^{-1} O_p \left( \frac{(\log n)^{1/2} + \delta_k^{-1} + A_k}{n^{1/2}} \right). \]

Since this last expression is independent of \( s \), this proves the third line of (S.3). The third line of (S.4) is proven in a similar manner.

**Lemma 4.** Assume (A1), (T1) and (T2) hold. Let \( A_k = \|\rho_k\|_\infty, M \) as in (A1), \( \delta_k \) as in (5.1), and \( C_1 \) as in (T1) with \( D_1 = M \). Let \( K^*(n) \to \infty \) be any sequence which satisfies \( \tau_{K^*} n^{1/2} \to \infty \) and
\[ \sum_{k=1}^{K^*} \left[ (\log n)^{1/2} + \delta_k^{-1} + A_k + \tau_{K^*} \delta_k^{-1} A_k \right] = O(\tau_{K^*} n^{1/2}). \]

Let \( C_2 \) be as in (T2), \( X_{i,K} = \nu + \sum_{k=1}^{K^*} \eta_{ik} \rho_k, \bar{X}_{i,K} = \tilde{\nu} + \sum_{k=1}^{K^*} \tilde{\eta}_{ik} \tilde{\rho}_k \), and set
\[ S_{K^*} = \max_{1 \leq K \leq K^*} \max_{1 \leq i \leq n} C_2(\|X_{i,K}\|_\infty, \|X'_{i,K}\|_\infty). \]

Then
\[ \max_{1 \leq K \leq K^*} \max_{1 \leq i \leq n} d(f_i(\cdot, K, \psi), \tilde{f}_i(\cdot, K, \psi)) = O_p \left( \frac{S_{K^*} \sum_{k=1}^{K^*} \delta_k^{-1}}{n^{1/2}} \right). \]

**Proof.** First, observe that \( f_i(\cdot, K, \psi) = \psi^{-1}(X_{i,K}) \) and \( \tilde{f}_i(\cdot, K, \psi) = \psi^{-1}(\bar{X}_{i,K}) \).

Recall that \( |\eta_{ik}| \leq 2C_1|T|^{1/2} \) for all \( i \) and \( k \) (see (4.13)). Then, by (A1) and Corollary 1,
\[ |\eta_{ik} - \tilde{\eta}_{ik}| \leq d_2(X_{i,K}, \nu) d_2(\rho_k, \tilde{\rho}_k) + d_2(\nu, \tilde{\nu}) = \delta_k^{-1} O_p(n^{-1/2}), \]
where the \( O_p \) term is uniform over \( i \) and \( k \). Next, using Lemma 3 and Corollary 1, along with the requirement that \( \tau_{K^*} n^{1/2} \to \infty \), for \( K \leq K^* \)
\[ d_\infty(X_{i,K}, \bar{X}_{i,K}) \leq d_\infty(\nu, \tilde{\nu}) + \sum_{k=1}^{K} d_\infty(\eta_{ik} \rho_k, \tilde{\eta}_{ik} \tilde{\rho}_k) \]
\[ \leq d_\infty(\nu, \tilde{\nu}) + \sum_{k=1}^{K} |\eta_{ik}| d_\infty(\rho_k, \tilde{\rho}_k) + \sum_{k=1}^{K} \|\rho_k\|_\infty |\eta_{ik} - \tilde{\eta}_{ik}| \]
\[ = O_p \left( \frac{\sum_{k=1}^{K} [(\log n)^{1/2} + \delta_k^{-1} + A_k + \tau_{K^*} \delta_k^{-1} A_k]}{\tau_{K^*} n^{1/2}} \right). \]
Since the $O_p$ term does not depend on $i$ or $K$, by the first assumption in the statement of the Lemma, we have
\[
\max_{1 \leq K \leq K^*} \max_{1 \leq i \leq n} d_\infty(X_{i,K}, \tilde{X}_{i,K}) = O_p(1).
\]

For $C_{3,K,i} = C_3(d_\infty(X_{i,K}, \tilde{X}_{i,K}))$ as in (T2),
\[
\max_{1 \leq K \leq K^*} \max_{1 \leq i \leq n} C_{3,K,i} = O_p(1),
\]
whence
\[
d_2(X_{i,K}, \tilde{X}_{i,K}) \leq d_2(\nu, \tilde{\nu}) + \sum_{k=1}^K d_2(\eta_{ik}\bar{\rho}_k, \tilde{\eta}_{ik}\tilde{\rho}_k)
\]
\[
\leq d_2(\nu, \tilde{\nu}) + \sum_{k=1}^K |\eta_{ik}d_2(\rho_k, \tilde{\rho}_k) + \sum_{k=1}^K |\eta_{ik} - \tilde{\eta}_{ik}|
\]
\[
= O_p\left(n^{-1/2} \sum_{k=1}^K \delta_k^{-1}\right).
\]

Again, this $O_p$ term does not depend on $i$ or $K$, so
\[
\max_{1 \leq K \leq K^*} \max_{1 \leq i \leq n} d_2(X_{i,K}, \tilde{X}_{i,K}) = O_p\left(n^{-1/2} \sum_{k=1}^{K^*} \delta_k^{-1}\right),
\]
leading to
\[
\max_{1 \leq K \leq K^*} \max_{1 \leq i \leq n} d(f_i(\cdot, K, \psi), \tilde{f}_i(\cdot, K, \psi)) \leq S_{K^*} \max_{1 \leq K \leq K^*} \max_{1 \leq i \leq n} C_{3,K,i} d_2(X_{i,K}, \tilde{X}_{i,K})
\]
\[
= O_p\left(S_{K^*} \sum_{k=1}^{K^*} \delta_k^{-1} \frac{1}{n^{1/2}}\right).
\]
\[\Box\]