Biometrika (xxxx), **xxx**, x, *pp*. 1–8
Printed in Great Britain

Supplementary Material for "Additive Models for Symmetric Positive-Definite Matrices and Lie Groups"

BY Z. LIN

Department of Statistics and Applied Probability, National University of Singapore
21 Lower Kent Ridge Road, 119077, Singapore
linz@nus.edu.sg

H.-G. MÜLLER

Department of Statistics, University of California, Davis
One Shields Avenue, Davis, CA 95616, U.S.A.
hgmueller@ucdavis.edu

AND B. U. PARK

Department of Statistics, Seoul National University
1 Gwanak-ro, Gwanak-gu, Seoul, Republic of Korea
bupark@stats.snu.ac.kr

FOR THE ALZHEIMER'S DISEASE NEUROIMAGING INITIATIVE*

S1. SOME COMPUTATIONAL DETAILS

For the algorithm in Section 3, one needs to compute the sample Fréchet mean in Step 1 and the Riemannian logarithmic map in Step 2. Below are the closed-form expressions for these quantities.

Log-Cholesky. According to Lin (2019), the Fréchet mean $\hat{\mu}$ of $Y_1, \ldots, Y_n$ is given by $\hat{L}\hat{L}^\top$ with $\hat{L} = n^{-1}\sum_{i=1}^n \mathfrak{L}(L_i) + \exp\left\{n^{-1}\sum_{i=1}^n \log \mathfrak{D}(L_i)\right\}$, where $L_1, \ldots, L_n$ are respectively the Cholesky factors of $Y_1, \ldots, Y_n$, $\mathfrak{L}(L_i)$ denotes the strict lower triangular part of L_i, and $\mathfrak{D}(L_i)$ denotes the diagonal part of L_i. For $S, Y \in \mathcal{S}_m^+$ respectively with Cholesky factors L and K, the Riemannian logarithmic map is given by $\mathrm{Log}_S Y = LZ^\top + ZL^\top$ with $Z = \mathfrak{L}(K) - \mathfrak{L}(L) + \mathfrak{D}(L)\log\{\mathfrak{D}(L)^{-1}\mathfrak{D}(K)\}$.

Log-Euclidean. According to Arsigny et al. (2007), the sample Fréchet mean of $Y_1, \ldots, Y_n \in \mathcal{S}_m^+$ when $\mathcal{S}_m^+$ is endowed with the Log-Euclidean metric is given by $\hat{\mu} = \exp\left(\frac{1}{n}\sum_{i=1}^n \log Y_i\right)$. In addition, $\mathrm{Log}_S Y = (D_{\log S}\exp)(\log Y - \log S)$ for any $S, Y \in \mathcal{S}_m^+$. According to Fillard et al. (2005) or Pennec (2020), for $S \in \mathcal{S}_m^+$ with the spectral decomposition $S = Q\mathrm{diag}(\lambda_1, \ldots, \lambda_m)Q^\top$ for a unitary matrix Q, for any $V \in \mathcal{S}_m$, we have $(D_S\exp)(V) = Q(W \odot (Q^\top V Q))Q^\top$, where W is an $m \times m$ matrix with elements $W_{jk} = \frac{e^{\lambda_j} - e^{\lambda_k}}{\lambda_j - \lambda_k}$ when $\lambda_j \neq \lambda_k$ and $W_{jk} = e^{\lambda_j}$ when $\lambda_j = \lambda_k$.

* Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (`adni.loni.usc.edu`). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: `http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf`

© 2020 Biometrika Trust

Table S1. *Prediction RMSE and its Monte Carlo standard error with* $m = 10$ *(Log-Cholesky)*

Setting	q	n	MAM		ILPR		CHOL	
			SNR=2	SNR=4	SNR=2	SNR=4	SNR=2	SNR=4
I	3	50	1.481 (0.171)	1.433 (0.139)	2.629 (0.169)	2.580 (0.239)	2.159 (0.115)	1.960 (0.145)
		100	1.052 (0.059)	0.938 (0.049)	2.532 (0.143)	2.485 (0.200)	1.892 (0.095)	1.815 (0.076)
		200	0.670 (0.032)	0.540 (0.032)	2.431 (0.186)	2.275 (0.375)	1.769 (0.069)	1.796 (0.070)
	4	50	2.447 (0.154)	2.430 (0.194)	3.243 (0.175)	3.188 (0.162)	3.148 (0.131)	3.079 (0.103)
		100	1.569 (0.112)	1.410 (0.068)	2.944 (0.120)	2.940 (0.140)	2.714 (0.113)	2.553 (0.063)
		200	0.991 (0.070)	0.821 (0.028)	2.884 (0.170)	2.854 (0.119)	2.662 (0.077)	2.564 (0.071)
II	3	50	1.852 (0.045)	1.829 (0.061)	2.223 (0.169)	2.189 (0.183)	1.965 (0.066)	1.926 (0.071)
		100	1.470 (0.070)	1.421 (0.029)	2.110 (0.175)	2.001 (0.146)	1.663 (0.105)	1.571 (0.076)
		200	1.282 (0.023)	1.249 (0.024)	2.030 (0.241)	1.958 (0.123)	1.415 (0.042)	1.352 (0.040)
	4	50	2.580 (0.051)	2.529 (0.080)	2.825 (0.125)	2.762 (0.127)	2.830 (0.108)	2.790 (0.070)
		100	1.988 (0.110)	1.820 (0.107)	2.623 (0.135)	2.594 (0.125)	2.515 (0.099)	2.410 (0.066)
		200	1.466 (0.052)	1.395 (0.015)	2.575 (0.129)	2.543 (0.119)	2.381 (0.090)	2.283 (0.082)
III	3	50	0.903 (0.132)	0.924 (0.104)	1.145 (0.145)	1.084 (0.156)	0.875 (0.128)	0.866 (0.096)
		100	0.884 (0.129)	0.825 (0.107)	1.000 (0.132)	0.912 (0.127)	0.864 (0.112)	0.812 (0.099)
		200	0.872 (0.132)	0.804 (0.087)	0.959 (0.155)	0.880 (0.104)	0.881 (0.124)	0.804 (0.083)
	4	50	0.865 (0.131)	0.802 (0.084)	1.078 (0.187)	1.071 (0.154)	0.854 (0.120)	0.834 (0.086)
		100	0.837 (0.081)	0.819 (0.066)	0.898 (0.147)	0.900 (0.124)	0.837 (0.102)	0.820 (0.065)
		200	0.820 (0.090)	0.804 (0.108)	0.823 (0.130)	0.816 (0.110)	0.826 (0.094)	0.821 (0.105)

S2. ADDITIONAL SIMULATION RESULTS

The prediction RMSE and its Monte Carlo standard error for $m = 10$ is presented in Tables S1 and S2 respectively for the Log-Cholesky and Log-Euclidean metrics.

Figure S1 provides a graphical illustration of predicted responses at some locations from a randomly selected simulation replicate, where the locations are chosen in the way that the squared errors of the MAM estimator at the locations are close to those in Table 2 under the setting of SNR $= 4$, $q = 3$ and $m = 3$, in order to visualize the estimation quality of the corresponding prediction RMSE.

The average computation time of each method for one simulation replicate is presented in Table S3; the computation was performed on a laptop with an Intel(R) Core(TM) i7-8650U CPU (1.90GHz) and 16G system memory. One finds that the Log-Cholesky metric is computationally more efficient than the Log-Euclidean metric. In addition, the proposed method with the Log-Cholesky metric and the baseline method have comparable efficiency, while the ILPR is computationally more efficient.

S3. PROOFS

Proof of Proposition 1. According to Lemma S1, $\mathrm{Log}_\mu(\mu \oplus z) = \tau_{e,\mu}\mathfrak{log}(z)$. Applying this with $z = w_1(X_1) \oplus \cdots \oplus w_q(X_q) \oplus \zeta$, we have $\mathrm{Log}_\mu Y = \tau_{e,\mu}\mathfrak{log}(w_1(X_1) \oplus \cdots \oplus w_q(X_q) \oplus \zeta) = \sum_{k=1}^q \tau_{e,\mu}\mathfrak{log} w_k(X_k) + \tau_{e,\mu}\mathfrak{log}\zeta$, where the second equality stems from $\mathfrak{exp}(\mathfrak{u} + \mathfrak{v}) = \mathfrak{exp}(\mathfrak{u}) \oplus \mathfrak{exp}(\mathfrak{v})$ for $\mathfrak{u}, \mathfrak{v} \in \mathfrak{g}$ and this leads to $\mathfrak{log}(u \oplus v) = \mathfrak{log}(u) + \mathfrak{log}(v)$ for $u, v \in \mathcal{L}$. □

Proof of Proposition 2. We utilize Corollary 1 of Schötz (2019). We first observe that $d(\hat{\mu}, \mu) = o_P(1)$ according to Theorem 2.3 of Bhattacharya & Patrangenaru (2003) and condition (A1). Then according to Schötz (2019) the growth and entropy conditions are required to hold only in a neighborhood of μ, where the corresponding existence and growth conditions are in assumptions (A1) and (A4), respectively. Condition (A1) implies that F is finite for some point and thus by the triangle inequality for all points in the manifold. If Z is an independent copy of Y, then $\mathbb{E}d^2(Y, Z) \leq 2\mathbb{E}\{d^2(Y, \mu)\} + 2\mathbb{E}\{d^2(\mu, Z)\} = 4F(\mu) < \infty$, and the moment

Table S2. *Prediction RMSE and its Monte Carlo standard error with $m = 10$ (Log-Euclidean)*

Setting	q	n	MAM		ILPR		CHOL	
			SNR=2	SNR=4	SNR=2	SNR=4	SNR=2	SNR=4
I	3	50	3.188 (0.313)	2.937 (0.395)	5.144 (0.481)	5.035 (0.222)	4.649 (0.229)	4.514 (0.267)
		100	2.093 (0.152)	1.909 (0.153)	4.936 (0.336)	4.912 (0.348)	4.424 (0.346)	4.344 (0.225)
		200	1.380 (0.042)	1.109 (0.059)	4.615 (0.306)	4.508 (0.288)	4.328 (0.149)	4.271 (0.305)
	4	50	5.495 (0.337)	4.948 (0.331)	6.409 (0.216)	6.353 (0.237)	6.895 (0.240)	6.548 (0.274)
		100	3.325 (0.237)	2.872 (0.294)	6.017 (0.182)	5.960 (0.157)	6.227 (0.279)	6.016 (0.272)
		200	2.058 (0.095)	1.750 (0.092)	5.841 (0.133)	5.758 (0.117)	6.005 (0.236)	5.919 (0.340)
II	3	50	3.895 (0.172)	3.779 (0.185)	4.766 (0.107)	4.707 (0.191)	4.330 (0.251)	4.156 (0.194)
		100	3.085 (0.085)	2.932 (0.094)	4.367 (0.240)	4.312 (0.204)	3.515 (0.141)	3.307 (0.132)
		200	2.716 (0.063)	2.636 (0.038)	4.112 (0.189)	4.023 (0.197)	3.127 (0.063)	3.032 (0.088)
	4	50	5.422 (0.125)	5.395 (0.144)	5.925 (0.142)	5.829 (0.061)	5.990 (0.155)	5.953 (0.170)
		100	4.254 (0.085)	3.917 (0.170)	5.410 (0.167)	5.353 (0.146)	5.315 (0.210)	5.202 (0.165)
		200	3.224 (0.073)	3.007 (0.078)	5.389 (0.149)	5.279 (0.126)	5.177 (0.191)	5.108 (0.331)
III	3	50	1.017 (0.116)	1.032 (0.110)	1.280 (0.143)	1.030 (0.129)	1.085 (0.184)	1.097 (0.203)
		100	0.913 (0.127)	0.926 (0.093)	1.172 (0.125)	0.954 (0.137)	0.981 (0.169)	0.983 (0.150)
		200	0.815 (0.085)	0.864 (0.073)	1.051 (0.103)	0.915 (0.114)	0.978 (0.202)	0.971 (0.112)
	4	50	1.047 (0.109)	0.999 (0.081)	1.095 (0.168)	1.009 (0.076)	1.074 (0.089)	0.993 (0.050)
		100	0.948 (0.088)	0.933 (0.092)	0.966 (0.150)	0.965 (0.155)	0.945 (0.095)	0.912 (0.091)
		200	0.939 (0.072)	0.917 (0.080)	0.947 (0.108)	0.924 (0.095)	0.908 (0.092)	0.882 (0.156)

Table S3. *Empirical computation time and its Monte Carlo standard error (in seconds)*

m	q	n	MAM		ILPR		CHOL
			Log-Cholesky	Log-Euclidean	Log-Cholesky	Log-Euclidean	
3	3	50	3.765 (0.091)	7.559 (0.214)	2.459 (0.138)	2.820 (0.071)	3.340 (0.157)
		100	4.880 (0.221)	9.854 (0.479)	2.784 (0.091)	3.329 (0.196)	4.691 (0.289)
		200	9.314 (0.753)	15.028 (0.671)	2.961 (0.282)	3.801 (0.120)	7.682 (0.408)
	4	50	9.112 (1.767)	13.362 (1.788)	6.471 (0.524)	6.915 (0.517)	8.795 (0.948)
		100	14.069 (1.136)	18.831 (1.066)	6.939 (0.574)	7.016 (0.212)	12.148 (0.707)
		200	19.161 (1.706)	24.549 (0.721)	8.407 (0.267)	9.221 (0.177)	18.490 (0.488)
10	3	50	26.689 (0.483)	33.633 (0.341)	26.044 (0.356)	28.671 (1.600)	21.797 (0.273)
		100	39.470 (2.828)	61.653 (2.715)	26.289 (3.004)	30.950 (2.530)	39.119 (2.502)
		200	101.813 (2.996)	115.781 (3.229)	27.930 (1.849)	33.179 (0.881)	97.222 (2.599)
	4	50	64.362 (0.786)	66.677 (0.923)	52.643 (0.369)	53.201 (0.266)	55.641 (1.460)
		100	95.829 (3.127)	128.186 (3.787)	60.457 (3.088)	63.669 (2.315)	85.028 (2.134)
		200	138.137 (0.304)	155.520 (3.325)	67.039 (0.943)	67.117 (1.762)	143.422 (1.859)

condition of Schötz (2019) follows, as well as the weak quadruple condition, where the latter holds for all Hadamard spaces and bounded spaces. Since a compact manifold is a bounded space, the weak quadruple condition holds for manifolds under (M1) or (M2).

Finally we verify the entropy condition of Schötz (2019). If $\mathcal{M}$ is compact, then its sectional curvature is bounded away from $-\infty$ and $+\infty$. According to the Bishop–Günther inequality (Gray, 2004, Eq. (3.34)), $\mathrm{vol}(B_\epsilon(z)) \geq C_1 \epsilon^D$ for all sufficiently small $\epsilon > 0$, where $B_\epsilon(z) = \{y \in \mathcal{M} : d(y, z) < \epsilon\}$ and C_1 is a constant independent of z and ϵ. With this result and equation (3.33) of Gray (2004), the packing number and thus the covering number of $B_\delta(\mu)$ is bounded by $O(\delta^D \epsilon^{-D})$. Therefore, the entropy condition holds for $\alpha = \beta$ for a sufficiently small neighborhood of μ and the result follows. □

The following lemmas are instrumental to establish Theorems 1 and 2.

LEMMA S1. *If $(\mathcal{L}, \oplus)$ is an abelian Lie group endowed with a bi-invariant metric that turns $\mathcal{L}$ into a Hadamard space, then for all $y, z \in \mathcal{M}$, $\mathrm{Log}_y(y \oplus z) = \tau_{e,y}\mathfrak{log}(z)$, where e is the group identity.*

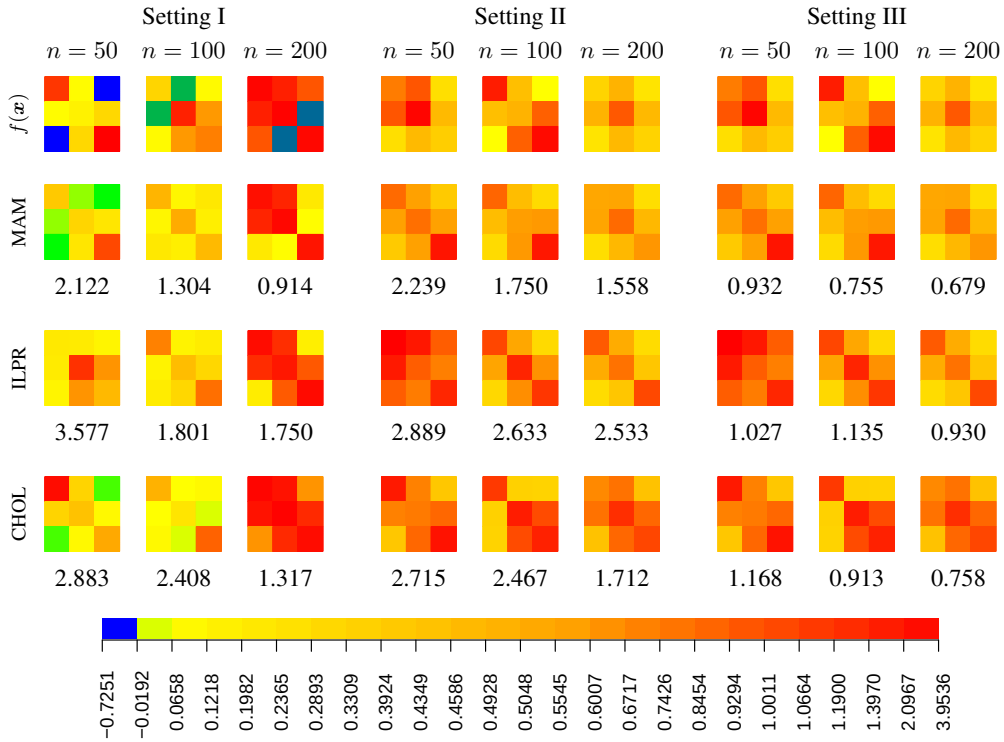


Fig. S1. Examples of predicted responses at some locations from a random simulation replicate such that the squared errors of the MAM estimator at these locations are close to those in Table 2, under the setting of SNR $= 4$, $q = 3$ and $m = 3$. The first row represents the target Fréchet means at those locations, while the next three rows represent the estimates obtained by the MAM, ILPR and CHOL methods. The numbers under the matrices are the corresponding squared errors.

*Proof of Lemma S*1. First, since $\mathcal{L}$ is a Hadamard manifold, the Riemannian logarithmic map Log_μ and the Lie logarithmic map $\mathfrak{log}$ are well defined for all elements of $\mathcal{L}$. Moreover, for a bi-invariant Lie group, the Riemannian exponential map Exp_e at the identity element e coincides with the Lie exponential map $\mathfrak{exp}$.

For $z \in \mathcal{L}$, let $\mathfrak{u} = \mathfrak{log}(z) \in \mathfrak{g}$ and denote by U the associated left-invariant vector field. Define $\gamma_y(t) = L_y(\mathfrak{exp}(t\mathfrak{u}))$. Then $\gamma_y'(t) = U(y \oplus \gamma_e(t))$ based on the proof of Lemma 6 in Lin (2019). The fact that $\gamma_e(0) = e$ further leads to $\gamma_y'(0) = U(y) = \tau_{e,y}\mathfrak{u}$, where the second equality is due to the fact that the parallel transport of $\mathfrak{u}$ is realized by the vector field U. Noting that $\gamma_y(0) = y$, by the definition of the Riemannian exponential map, $\mathrm{Exp}_y\gamma_y'(0) = \gamma_y(1)$, which leads to $\mathrm{Exp}_y\tau_{e,y}\mathfrak{u} = L_y(\mathfrak{exp}(\mathfrak{u})) = y \oplus z$. The last equation is equivalent to $\mathrm{Log}_y(y \oplus z) = \tau_{e,y}\mathfrak{log}(z)$. □

LEMMA S2. *If $Z_1, \ldots, Z_n$ are nonnegative i.i.d. random variables with $\mathbb{E}Z_1^\alpha < \infty$ for some $\alpha > 0$, then* $\max_{1 \le i \le n} Z_i = O_P(n^{1/\alpha})$.

Proof of Lemma S2. Let $a_n = n^{1/\alpha}$. By i.i.d. assumption, for $\epsilon > 0$,

$$\begin{aligned}\Pr\left\{\max_{1\le i\le n} Z_i \le Ca_n\right\} &= (\Pr\{Z_1 \le Ca_n)^n = (1-\Pr\{Z_1 > Ca_n\})^n \\ &\ge \left(1-\frac{\mathbb{E}Z_1^\alpha}{C^\alpha a_n^\alpha}\right)^n = \left(1-\frac{\mathbb{E}Z_1^\alpha}{C^\alpha n}\right)^n \to e^{-\mathbb{E}Z_1^\alpha/C^\alpha} \\ &\ge 1-\epsilon\end{aligned}$$

for a sufficiently large C that depends on $\epsilon > 0$. □

LEMMA S3. *Assume the conditions (A1)–(A5) and (B4). If $\mathbb{E}\|\varepsilon\|_\mu^\alpha < \infty$ for some $\alpha > 2$, then*

$$\max_{1\le i\le n} \|\tau_{\hat\mu,\mu}\mathrm{Log}_{\hat\mu} Y_i - \mathrm{Log}_\mu Y_i\|_\mu = O_P(n^{-(\alpha-2)/(2\alpha)}).$$

Proof of Lemma S3. Using the inequality (5.7) of Kendall & Le (2011), the condition (A5) and $d(\hat\mu,\mu) = o_P(1)$ that is guaranteed by Proposition 2, we deduce that, with probability tending to one,

$$\max_{1\le i\le n} \|\tau_{\hat\mu,\mu}\mathrm{Log}_{\hat\mu} Y_i - \mathrm{Log}_\mu Y_i\|_\mu = O(1)d(\hat\mu,\mu)\max_{1\le i\le n}\|\mathrm{Log}_\mu Y_i\|_\mu.$$

By Lemma S2, the moment condition $\mathbb{E}\|\varepsilon\|_\mu^\alpha < \infty$, the compactness of $\mathcal{X}$ and the continuity of $f_1,\dots,f_q$ assumed in (B4), we have $\max_{1\le i\le n}\|\mathrm{Log}_\mu Y_i\|_\mu = O_P(n^{1/\alpha})$. The conclusion of the lemma then follows from Proposition 2. □

Proof of Theorem 1. Define $\tilde m_j$ as $\hat m_j$ in (4) with $\mathrm{Log}_{\hat\mu} Y_i$ being replaced by $\mathrm{Log}_\mu Y_i$. Let $(\tilde f_j : 1 \le j \le q)$ denote the solution of the system of equations

$$\tilde f_k(x_k) = \tilde m_k(x_k) - n^{-1}\sum_{i=1}^n \mathrm{Log}_\mu Y_i - \sum_{j:j\ne k}\int_{\mathcal{X}_j} \tilde f_j(x_j)\frac{\hat p_{kj}(x_k,x_j)}{\hat p_k(x_k)}\mathrm{d}x_j, \quad 1\le k\le q,$$

subject to the constraints $\int_{\mathcal{X}_k}\tilde f_k(x_k)\hat p_k(x_k)\mathrm{d}x_k = 0$ for $1\le k\le q$. According to the theory of Jeon & Park (2020), under the conditions of the theorem, the solution exists and is unique with probability tending to one. Furthermore, it holds that

$$\begin{aligned}\max_{1\le k\le q}\int_{\mathcal{I}_k}\|\tilde f_k(x_k) - f_k(x_k)\|_\mu^2\, p_k(x_k)\mathrm{d}x_k &= O_P(n^{-4/5}), \\ \max_{1\le k\le q}\int_{\mathcal{X}_k}\|\tilde f_k(x_k) - f_k(x_k)\|_\mu^2\, p_k(x_k)\mathrm{d}x_k &= O_P(n^{-3/5}).\end{aligned} \tag{S1}$$

Since the smooth backfitting operation at (3) is linear in the response variables and the parallel transport $\tau_{\hat\mu,\mu} : T_{\hat\mu}(\mathcal{M}) \to T_\mu(\mathcal{M})$ is also a linear map, we get that $(\tau_{\hat\mu,\mu}\hat f_j : 1\le j\le q)$ is nothing else than the smooth backfitting estimator that one gets from the smooth backfitting operation with $\tau_{\hat\mu,\mu}\mathrm{Log}_{\hat\mu} Y_i$ as responses. If we can show

$$\max_{1\le k\le q}\int_{\mathcal{X}_k}\|\tau_{\hat\mu,\mu}\hat f_k(x_k) - \tilde f_k(x_k)\|_\mu^2\, p_k(x_k)\mathrm{d}x_k = O_P(n^{-(\alpha-2)/\alpha}), \tag{S2}$$

then (S1) and (S2) imply the theorem.

To prove the claim (S2), let $\delta_k = \tau_{\hat{\mu},\mu}\hat{f}_k - \tilde{f}_k$. Then, $(\delta_j : 1 \le j \le q)$ is the solution of the system of equations

$$\delta_k(x_k) = \tau_{\hat{\mu},\mu}\hat{m}_k(x_k) - \tilde{m}_k(x_k) - n^{-1}\sum_{i=1}^{n}\left(\tau_{\hat{\mu},\mu}\mathrm{Log}_{\hat{\mu}}Y_i - \mathrm{Log}_{\mu}Y_i\right) - \sum_{j:j\neq k}\int_{\mathcal{X}_j}\delta_j(x_j)\frac{\hat{p}_{kj}(x_k,x_j)}{\hat{p}_k(x_k)}\mathrm{d}x_j, \quad 1 \le k \le q,$$

subject to the constraints $\int_{\mathcal{X}_k}\delta_k(x_k)\hat{p}_k(x_k)\mathrm{d}x_k = 0$ for $1 \le k \le q$. From Lemma S3 it follows that

$$\left\|\tau_{\hat{\mu},\mu}\hat{m}_k(x_k) - \tilde{m}_k(x_k) - n^{-1}\sum_{i=1}^{n}\left(\tau_{\hat{\mu},\mu}\mathrm{Log}_{\hat{\mu}}Y_i - \mathrm{Log}_{\mu}Y_i\right)\right\|_{\mu}^2 = O_P(n^{-(\alpha-2)/\alpha})$$

uniformly for $x_k \in \mathcal{X}_k$ for all $1 \le k \le q$. Using the arguments in the proof of Theorem 4.1 in Jeon & Park (2020), we may then prove

$$\sup_{x_k\in\mathcal{X}_k}\|\delta_k(x_k)\|_{\mu}^2 = O_P(n^{-(\alpha-2)/\alpha}). \tag{S3}$$

This gives (S2). □

Proof of Theorem 2. Let $\tilde{m}_j^A(x_j) = n^{-1}\hat{p}_j(x_j)^{-1}\sum_{i=1}^{n}K_{h_j}(x_j,X_{ij})\varepsilon_i$. Then we find

$$n^{2/5}\left(\tilde{f}_j(x_j) - f_j(x_j)\right) = n^{2/5}\tilde{m}_j^A(x_j) + \tau_{e,\mu}\left(\frac{1}{2}\alpha_j^2 u_2\cdot\psi_j''(x_j) + \Delta_j(x_j)\right) + o_P(1) \tag{S4}$$

for each $x_j \in \mathrm{Int}(\mathcal{X})$ for all $1 \le j \le q$, where $u_2 = \int u^2K(u)\mathrm{d}u$ and $\tilde{f}_j$ is defined in the proof of Theorem 1. Here, $W_n = o_P(1)$ means $\lim_{n\to\infty}P(\|W\|_{\mu} > \epsilon) = 0$ for all $\epsilon > 0$. The assertion (S4) can be proved along the lines of the proof of Theorem 4.3 in Jeon & Park (2020). The expansion (S4) together with (S3) entails

$$n^{2/5}\left(\tau_{\hat{\mu},\mu}\hat{f}_j(x_j) - f_j(x_j)\right) = n^{-3/5}\hat{p}_j(x_j)^{-1}\sum_{i=1}^{n}K_{h_j}(x_j,X_{ij})\varepsilon_i + \theta_j(x_j) + o_P(1), 1 \le j \le q.$$

Here we use $\alpha > 10$. By identifying $T_{\mu}\mathcal{M}$ and its metric g_{μ} with the Hilbert space $\mathbb{H}$ and the associated inner product $\langle\cdot,\cdot\rangle$ in Jeon & Park (2020), respectively, and utilizing Theorem 1.1 in Kundu et al. (2000), one finds that the joint distribution of $\left(n^{-3/5}\hat{p}_j(x_j)^{-1}\sum_{i=1}^{n}K_{h_j}(x_j,X_{ij})\varepsilon_i : 1 \le j \le q\right)$ converges to $N(\boldsymbol{x})$. This completes the proof of the theorem. □

REFERENCES

ARSIGNY, V., FILLARD, P., PENNEC, X. & AYACHE, N. (2007). Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal of Matrix Analysis and Applications* **29**, 328–347.

BHATTACHARYA, R. & PATRANGENARU, V. (2003). Large sample theory of intrinsic and extrinsic sample means on manifolds. I. *The Annals of Statistics* **31**, 1–29.

FILLARD, P., ARSIGNY, V., PENNEC, X. & AYACHE, N. (2005). Joint estimation and smoothing of clinical dt-mri with a log-euclidean metric. Tech. Rep. RR-5607, INRIA.

GRAY, A. (2004). *Tubes*. Springer Basel AG, 2nd ed.

JEON, J. M. & PARK, B. U. (2020). Additive regression with Hilbertian responses. *The Annals of Statistics* **48**, 2671–2697.

KENDALL, W. S. & LE, H. (2011). Limit theorems for empirical Fréchet means of independent and non-identically distributed manifold-valued random variables. *Brazilian Journal of Probability and Statistics* **25**, 323–352.

KUNDU, S., MAJUMDAR, S. & MUKHERJEE, K. (2000). Central limit theorems revisited. *Statistics and Probability Letters* **47**, 265–275.

LIN, Z. (2019). Riemannian geometry of symmetric positive definite matrices via Cholesky decomposition. *SIAM Journal on Matrix Analysis and Applications* **40**, 1353–1370.

PENNEC, X. (2020). Manifold-valued image processing with spd matrices. In *Riemannian Geometric Statistics in Medical Image Analysis*, X. Pennec, S. Sommer & T. Fletcher, eds., chap. 3. Elsevier, pp. 75–134.

SCHÖTZ, C. (2019). Convergence rates for the generalized Fréchet mean via the quadruple inequality. *Electronic Journal of Statistics* **13**, 4280–4345.

[*Received on* x xxx xxxx. *Editorial decision on* x xxx xxxx]