Reply to "On some problems in the article *Efficient likelihood estimation in state space models* by Cheng-Der Fuh (*Annals of Statistics,* 34, 2026-2068, 2006) by Jens Ledet Jensen

Cheng-Der Fuh
National Central University and Academia Sinica

The author is grateful for the comments by Dr. Jensen. This note replies his comments.

**Problem 2.1 Definition of iterated function system**

$$\mathbf{P}_\theta(\xi_j)h(x) = \int_{y\in\mathcal{X}} p_\theta(y,x) f(\xi_j;\theta|x,\xi_{j-1})h(y)m(dy). \tag{2.6}$$

Define the composition of two random functions as

$$\begin{aligned} &\mathbf{P}_\theta(\xi_{j+1})\circ\mathbf{P}_\theta(\xi_j)h(x) \\ = &\int_{z\in\mathcal{X}} p_\theta(z,x)f(\xi_{j+1};\theta|x,\xi_j)\Big(\int_{y\in\mathcal{X}} p_\theta(y,z)f(\xi_j;\theta|z,\xi_{j-1})h(y)m(dy)\Big)m(dz). \end{aligned} \tag{2.7}$$

Page 2042. C1. $\cdots$ for all $s_0, s_1 \in \mathbf{R}^d$, and $\sup_{x\in\mathcal{X}}\int p_\theta(y,x)m(dy) < \infty$. Since $m$ is $\sigma$-finite, there exist pairwise disjoint $\mathcal{X}_n$ such that $\mathcal{X} = \cup_{n=1}^\infty \mathcal{X}_n$, and $0 < m(\mathcal{X}_n) < \infty$. Assume $E[\sum_{n=1}^\infty \frac{1}{2^n}\sup_{x\in\mathcal{X}_n} f(\xi_1;\theta|x,s_0)] < \infty$ for all $s_0 \in \mathbf{R}^d$. Denote $g_\theta(\xi_0,\xi_1) = \sup_{x\in\mathcal{X}}\int p_\theta(y,x)f(\xi_1;\theta|x,\xi_0)m(dy)$. Furthermore, we assume that there exists $p \geq 1$ as in K2 such that

$$\sup_{(x_0,s_0)\in\mathcal{X}\times\mathbf{R}^d} E^\theta_{(x_0,s_0)}\Big\{\log\Big(g_\theta(s_0,\xi_1)\cdots g_\theta(\xi_{p-1},\xi_p)\frac{w(X_p,\xi_p)}{w(x_0,s_0)}\Big)\Big\} < 0, \tag{5.2}$$

The example on Page 2044, L12 hold if $\alpha \neq 0$. The original (5.6) was wrong, It should be

$$M_n := \mathbf{P}_\theta(\xi_n)\circ\cdots\circ\mathbf{P}_\theta(\xi_1)\circ\mathbf{P}_\theta(\xi_0)\pi \qquad \text{(Page 2045)} \tag{5.6}$$

Page 2046. LEMMA 3. $\cdots$. *Furthermore, under conditions C1, C6-C9, the function $g$ defined in (5.7) belongs to $\mathcal{L}(Q\times Q)$.*

PROOF OF LEMMA 3. We consider only the case of $\mathbf{P}(\xi_1)$, since the case of $\mathbf{P}(\xi_0)$ and $\mathbf{P}(\xi_j)$, for $j = 2,\cdots,n$, is a straightforward consequence. For any two elements $h_1, h_2 \in \mathbf{M}$, and two fixed elements $s_0, s_1 \in \mathbf{R}^d$, by (5.8) we have

$$\begin{aligned} &d(\mathbf{P}(s_1)h_1, \mathbf{P}(s_1)h_2) \\ = &\sup_{x\in\mathcal{X}}\Big|\int p_\theta(y,x)f(s_1;\theta|x,s_0)h_1(y)m(dy) - \int p_\theta(y,x)f(s_1;\theta|x,s_0)h_2(y)m(dy)\Big| \\ \leq &d(h_1,h_2)\sup_{x\in\mathcal{X}}\int p_\theta(y,x)f(s_1;\theta|x,s_0)m(dy) \leq C\Big(\sup_{x\in\mathcal{X}}\int p_\theta(y,x)m(dy)\Big)d(h_1,h_2), \end{aligned}$$

where $0 < C = \sup_{x\in\mathcal{X}} f(s_1;\theta|x,s_0) < \infty$, by assumption C1, is a constant. Note that $\sup_{x\in\mathcal{X}} \int p_\theta(y,x)m(dy) < \infty$ by assumption C1. The equality hold only if $h_1 = h_2$ $m$-almost surely. This proves the condition of Lipschitz continuous in the second argument.

Note that C1 implies K1 hold. Recall that $M_n = \mathbf{P}(\xi_n)\circ\cdots\circ\mathbf{P}(\xi_1)\circ\mathbf{P}(\xi_0)\pi$ for $\pi\in\mathbf{M}$ in (5.6). To prove the weighted mean contraction property K2, we observe that for $p \geq 1$,

$$
\begin{aligned}
(7.1)\ & \sup_{x_0,s_0} \mathbf{E}_{(x_0,s_0)}\Big\{\log\Big(L_p\frac{w(X_p,\xi_p)}{w(x_0,s_0)}\Big)\Big\} \\
= & \sup_{x_0,s_0} \mathbf{E}_{(x_0,s_0)}\Big\{\log\Big(\sup_{h_1\neq h_2}\frac{d(M_ph_1,M_ph_2)}{d(h_1,h_2)}\frac{w(X_p,\xi_p)}{w(x_0,s_0)}\Big)\Big\} \\
< & \sup_{x_0,s_0} \mathbf{E}_{(x_0,s_0)}\Big\{\log\Big(\prod_{j=1}^{p}\Big[\sup_{x_j\in\mathcal{X}}\int p_\theta(x_{j-1},x_j)f(\xi_j;\theta|x_j,s_{j-1})m(dx_{j-1})\Big]\frac{w(X_p,\xi_p)}{w(x_0,s_0)}\Big)\Big\} < 0.
\end{aligned}
$$

The last inequality follows from (5.2) in condition C1.

To verify assumption K3 hold. As $m$ is $\sigma$-finite, we have $\mathcal{X} = \cup_{n=1}^{\infty}\mathcal{X}_n$, where the $\mathcal{X}_n$ are pairwise disjoint and $0 < m(\mathcal{X}_n) < \infty$. Set

$$
h(x) = \sum_{n=1}^{\infty}\frac{I_{\mathcal{X}_n}(x)}{2^n m(\mathcal{X}_n)}. \tag{7.2}
$$

It is ease to see that $\int_{x\in\mathcal{X}} h(x)m(dx) = 1$ and hence belongs to $\mathbf{M}$. Observe that

$$
\begin{aligned}
(7.3)\qquad & \mathbf{E}d^2(\mathbf{P}(\xi_1)h,h) \\
= & \mathbf{E}\Big[\sup_{x_1\in\mathcal{X}}|\int p_\theta(x_0,x_1)f(\xi_1;\theta|x_1,s_0)h(x_0)m(dx_0) - h(x_1)|\Big] \\
\leq & \mathbf{E}\Big[\sum_{n=1}^{\infty}\frac{1}{2^n}\sup_{x_1\in\mathcal{X}_n} f(\xi_1;\theta|x_1,s_0)\Big]\Big[\sup_{x_1\in\mathcal{X}}\int p_\theta(x_0,x_1)m(dx_0)\Big] + \sup_{x_1\in\mathcal{X}}|h(x_1)|.
\end{aligned}
$$

Note that $h(x)$ is piecewise constant by definition (7.2), $E[\sum_{n=1}^{\infty}\frac{1}{2^n}\sup_{x\in\mathcal{X}_n} f(\xi_1;\theta|x,s_0)] < \infty$ for all $s_0\in\mathbf{R}^n$ by assumption C1, and $p_\theta(x_0,x_1)$ is integrable of $x_0$ over the subset $\mathcal{X}_n$ by assumption C1. These imply (7.3) is finite.

Finally, we observe

$$
\begin{aligned}
& \sup_{x_0,s_0}\mathbf{E}_{(x_0,s_0)}\Big\{L_1\frac{w(X_1,\xi_1)}{w(x_0,s_0)}\Big\} = \sup_{x_0,s_0}\mathbf{E}_{(x_0,s_0)}\Big\{\sup_{h_1\neq h_2}\frac{d(\mathbf{P}(\xi_1)h_1,\mathbf{P}(\xi_1)h_2)}{d(h_1,h_2)}\frac{w(X_1,\xi_1)}{w(x_0,s_0)}\Big\} \\
< & \sup_{x_0,s_0}\mathbf{E}_{(x_0,s_0)}\Big\{\Big(\sup_{x_1\in\mathcal{X}}\int p_\theta(x_0,x_1)f(\xi_1;\theta|x_1,s_0)m(dx_0)\Big)\frac{w(X_1,\xi_1)}{w(x_0,s_0)}\Big\} < \infty.
\end{aligned}
$$

The last inequality follows from (5.3) in condition C1.

Note that C8-C9 imply that $g\in\mathcal{L}(Q\times Q)$. Hence, the proof is completed. $\square$

**Problem 2.2 Harris recurrence of iterated function**

This paper is an extension of Fuh (2003) for finite state space, in which the likelihood function can be expressed as the $L_1$-norm of products of Markovian random matrices. Note that $M_n$ defined in (5.6) is an iterated random functions system govern by a Markov chain $Y_n$. And $Y_n = (X_n, \xi_n)$ in the state space models case. In Theorem 1 I only assume $Y_n = (X_n, \xi_n)$ is Harris recurrent. The purpose of the statement *"Note that under K1-K3, $\cdots$ a Markovian iterated random functions system in Theorem 2"* is to relate Theorem 1 and Theorem 2, to which I can apply limiting theorems in Markov chains to the law of large numbers and central limit theorem (and Edgeworth expansion) for $(Y_n, M_n)$.

In Lemma 4 I want to prove $Z_n = ((X_n, \xi_n), M_n)$ is Harris recurrent ($Z_n$ is defined in lines 1-2 on page 2056.) In the proof, I can use the results in Theorem 1 since only $Y_n = (X_n, \xi_n)$ is assumed to be Harris recurrent in Theorem 1. It is known that C1 implies $Y_n = (X_n, \xi_n)$ is Harris recurrent. A new proof of Lemma 3 was given on pages 1-2.

**Problem 2.3 Asymptotic properties of score function and observed information**

Page 2060, L12. In the proof of Lemma 6, (7.9) defined a new iterated functions system, therefore Corollary 1 can not be used directly. The same situation happens for Theorems 5 and 7. The rigorous proofs of these results will be given in a separate paper.

**Problem 2.4 Generality of conditions**

C5. For $\theta \in N_\delta(\theta_0)$,

$$E_x^\theta \Big( \frac{\partial \log \int_{y \in \mathcal{X}} \pi(x) p(x,y) f(s_0; \theta | x) f(\xi_1; \theta | y, s_0) m(dy)}{\partial \theta_i} \Big)^2 < w(x, s_0),$$

for all $i = 1, \cdots, q$.

Change C5' accordingly. It is straightforward to check that C5 holds for the Examples considered in Section 6. The proof of Lemma 5 can be done under C5.

**Other typos and mistakes.**

Page 2032, L1. $\cdots p_\theta(y,x) f(\xi_j; \theta | x, \xi_{j-1}) \cdots X_{j-1} = y$ and $X_j \in dx, \cdots$

$$\pi(y)\mathbf{P}(Y_n \in dz, M_n \in \cdot | Y_0 = y) = \pi(z)\tilde{\mathbf{P}}(\tilde{Y}_n \in dy, \tilde{M}_n \in \cdot | \tilde{Y}_0 = z) \tag{3.7}$$

Page 2028, L-5. $(1-\alpha^2)$. Page 2043. C7. $\theta \to \varphi_x(\theta)$ was a typo, delete it. Page 2047, L1. Then "each component of" the Fisher information matrix. L5. Replace "positive definite" by "finite". Page 2048, Theorem 5. Assume $\mathbf{I}(\theta_0)$ is invertible. Page 2057, L-3. The notation $m \times Q \times Q$ may be confusing, change it to $m \times Q \times \bar{Q}$.