Supplementary Material for "Efficient Maximum Approximated Likelihood Inference for Tukey's g-and-h Distribution"

Ganggang Xu[1] and Marc G. Genton[2]

1 Derivation of the Gradient and Hessian Matrix

Define the function $T(z_{p_{\boldsymbol\theta}(y)}, \boldsymbol\theta) = \xi + \omega\tau_{g,h}\{z_{p_{\boldsymbol\theta}(y)}\} = y$. For simplicity, from now on we just write $z_{p_{\boldsymbol\theta}(y)}$ as z_p. Then by the chain rule we have

$$\frac{\partial T(z_p,\boldsymbol\theta)}{\partial\boldsymbol\theta} = \frac{\partial T(z_p,\boldsymbol\theta)}{\partial^\dagger\boldsymbol\theta} + \frac{\partial T(z_p,\boldsymbol\theta)}{\partial^\dagger z_p}\frac{\partial z_p}{\partial\boldsymbol\theta} = 0,$$

and

$$\begin{aligned}\frac{\partial^2 T(z_p,\boldsymbol\theta)}{\partial\boldsymbol\theta\partial\boldsymbol\theta^T} = \frac{\partial^2 T(z_p,\boldsymbol\theta)}{\partial^\dagger\boldsymbol\theta\partial^\dagger\boldsymbol\theta^T} + \frac{\partial^2 T(z_p,\boldsymbol\theta)}{\partial^\dagger\boldsymbol\theta\partial^\dagger z_p}\frac{\partial z_p}{\partial\boldsymbol\theta^T} + \frac{\partial z_p}{\partial\boldsymbol\theta}\left\{\frac{\partial^2 T(z_p,\boldsymbol\theta)}{\partial^\dagger z_p\partial^\dagger\boldsymbol\theta^T} + \frac{\partial^2 T(z_p,\boldsymbol\theta)}{\partial^\dagger z_p^2}\frac{\partial z_p}{\partial\boldsymbol\theta^T}\right\}\\ + \frac{\partial T(z_p,\boldsymbol\theta)}{\partial^\dagger z_p}\frac{\partial^2 z_p}{\partial\boldsymbol\theta\partial\boldsymbol\theta^T} = 0,\end{aligned}$$

which implies

$$\frac{\partial z_p}{\partial\boldsymbol\theta} = -\frac{\partial T(z_p,\boldsymbol\theta)/\partial^\dagger\boldsymbol\theta}{\partial T(z_p,\boldsymbol\theta)/\partial^\dagger z_p},$$

and

$$\frac{\partial^2 z_p}{\partial\boldsymbol\theta\partial\boldsymbol\theta^T} = \left[\frac{\partial^2 T(z_p,\boldsymbol\theta)}{\partial^\dagger\boldsymbol\theta\partial^\dagger\boldsymbol\theta^T} + \frac{\partial^2 T(z_p,\boldsymbol\theta)}{\partial^\dagger\boldsymbol\theta\partial^\dagger z_p}\frac{\partial z_p}{\partial\boldsymbol\theta^T} + \frac{\partial z_p}{\partial\boldsymbol\theta}\left\{\frac{\partial^2 T(z_p,\boldsymbol\theta)}{\partial^\dagger z_p\partial^\dagger\boldsymbol\theta^T} + \frac{\partial^2 T(z_p,\boldsymbol\theta)}{\partial^\dagger z_p^2}\frac{\partial z_p}{\partial\boldsymbol\theta^T}\right\}\right]\left\{\frac{\partial T(z_p,\boldsymbol\theta)}{\partial^\dagger z_p}\right\}^{-1}.$$

[1]Department of Mathematical Sciences, Binghamton University, State University of New York, Binghamton, New York 13902, USA.
E-mail: gang@math.binghamton.edu

[2]CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia.
E-mail: marc.genton@kaust.edu.sa

As examples, we give the following quantities:

$$\begin{aligned}
\frac{\partial T(z_p,\boldsymbol{\theta})}{\partial^\dagger \xi} &= 1,\\
\frac{\partial T(z_p,\boldsymbol{\theta})}{\partial^\dagger \omega} &= g^{-1}\{\exp(gz_p)-1\}\exp(\frac{h}{2}z_p^2),\\
\frac{\partial T(z_p,\boldsymbol{\theta})}{\partial^\dagger g} &= -\omega g^{-2}\{\exp(gz_p)-1\}\exp(\frac{h}{2}z_p^2)+\omega g^{-1}\exp(gz_p+\frac{h}{2}z_p^2)z_p,\\
\frac{\partial T(z_p,\boldsymbol{\theta})}{\partial^\dagger h} &= \omega g^{-1}\{\exp(gz_p)-1\}\exp(\frac{h}{2}z_p^2)\frac{z_p^2}{2},\\
\frac{\partial T(z_p,\boldsymbol{\theta})}{\partial^\dagger z_p} &= \omega\exp(\frac{h}{2}z_p^2)[\exp(gz_p)+g^{-1}\{\exp(gz_p)-1\}hz_p],
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial^2 T(z_p,\boldsymbol{\theta})}{\partial^\dagger \xi^2} &= \frac{\partial^2 T(z_p,\boldsymbol{\theta})}{\partial^\dagger \xi\partial^\dagger\omega}=\frac{\partial^2 T(z_p,\boldsymbol{\theta})}{\partial^\dagger \xi\partial^\dagger g}=\frac{\partial^2 T(z_p,\boldsymbol{\theta})}{\partial^\dagger \xi\partial^\dagger h}=\frac{\partial^2 T(z_p,\boldsymbol{\theta})}{\partial^\dagger \xi\partial^\dagger z_p}=0,\\
\frac{\partial^2 T(z_p,\boldsymbol{\theta})}{\partial^\dagger \omega^2} &= 0,\\
\frac{\partial^2 T(z_p,\boldsymbol{\theta})}{\partial^\dagger \omega\partial^\dagger g} &= -g^{-2}\{\exp(gz_p)-1\}\exp(\frac{h}{2}z_p^2)+g^{-1}\exp(gz_p+\frac{h}{2}z_p^2)z_p,\\
\frac{\partial^2 T(z_p,\boldsymbol{\theta})}{\partial^\dagger \omega\partial^\dagger h} &= g^{-1}\{\exp(gz_p)-1\}\exp(\frac{h}{2}z_p^2)\frac{z_p^2}{2},\\
\frac{\partial^2 T(z_p,\boldsymbol{\theta})}{\partial^\dagger \omega\partial^\dagger z_p} &= \exp(\frac{h}{2}z_p^2)[\exp(gz_p)+g^{-1}\{\exp(gz_p)-1\}hz_p],\\
\frac{\partial^2 T(z_p,\boldsymbol{\theta})}{\partial^\dagger g^2} &= \omega\exp(\frac{h}{2}z_p^2)[2g^{-3}\{\exp(gz_p)-1\}-2g^{-2}\exp(gz_p)+g^{-1}\exp(gz_p)z_p^2],\\
\frac{\partial^2 T(z_p,\boldsymbol{\theta})}{\partial^\dagger g\partial^\dagger h} &= \frac{\omega}{2}\left\{-g^{-2}\{\exp(gz_p)-1\}+g^{-1}\exp(gz_p)z_p\right\}\exp(\frac{h}{2}z_p^2)z_p^2,\\
\frac{\partial^2 T(z_p,\boldsymbol{\theta})}{\partial^\dagger g\partial^\dagger z_p} &= \omega\exp(\frac{h}{2}z_p^2)\left[z_p\exp(gz_p)-g^{-2}\{\exp(gz_p)-1\}hz_p+g^{-1}\exp(gz_p)hz_p^2\}\right],\\
\frac{\partial^2 T(z_p,\boldsymbol{\theta})}{\partial^\dagger h^2} &= \frac{\omega}{4}g^{-1}\{\exp(gz_p)-1\}\exp(\frac{h}{2}z_p^2)z_p^4,\\
\frac{\partial^2 T(z_p,\boldsymbol{\theta})}{\partial^\dagger h\partial^\dagger z_p} &= \frac{\partial T(z_p,\boldsymbol{\theta})}{\partial^\dagger z_p}\frac{z_p^2}{2}+\omega\exp(\frac{h}{2}z_p^2)[g^{-1}\{\exp(gz_p)-1\}z_p].
\end{aligned}$$

Similarly, we can derive

$$\frac{\partial\varphi_{\boldsymbol{\theta}}(z_p)}{\partial\boldsymbol{\theta}}=\frac{\partial\varphi_{\boldsymbol{\theta}}(z_p)}{\partial^\dagger\boldsymbol{\theta}}+\frac{\partial\varphi_{\boldsymbol{\theta}}(z_p)}{\partial^\dagger z_p}\frac{\partial z_p}{\partial\boldsymbol{\theta}},$$

and

$$\frac{\partial^2 \varphi_{\boldsymbol{\theta}}(z_p)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} = \frac{\partial^2 \varphi_{\boldsymbol{\theta}}(z_p)}{\partial^\dagger \boldsymbol{\theta} \partial^\dagger \boldsymbol{\theta}^T} + \frac{\partial^2 \varphi_{\boldsymbol{\theta}}(z_p)}{\partial^\dagger \boldsymbol{\theta} \partial^\dagger z_p} \frac{\partial z_p}{\partial \boldsymbol{\theta}^T} + \frac{\partial z_p}{\partial \boldsymbol{\theta}} \left\{ \frac{\partial^2 \varphi_{\boldsymbol{\theta}}(z_p)}{\partial^\dagger z_p \partial^\dagger \boldsymbol{\theta}^T} + \frac{\partial^2 \varphi_{\boldsymbol{\theta}}(z_p)}{\partial^\dagger z_p^2} \frac{\partial z_p}{\partial \boldsymbol{\theta}^T} \right\}$$
$$+ \frac{\partial \varphi_{\boldsymbol{\theta}}(z_p)}{\partial^\dagger z_p} \frac{\partial^2 z_p}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}.$$

Finally, the Gradient and Hessian matrices of $L_n(\boldsymbol{\theta})$ can be calculated as

$$U_n(\boldsymbol{\theta}) = \sum_{i=1}^{n} \frac{\partial \varphi_{\boldsymbol{\theta}}(z_{p_i})}{\partial \boldsymbol{\theta}}, \tag{S.1}$$

and

$$I_n(\boldsymbol{\theta}) = \sum_{i=1}^{n} \frac{\partial^2 \varphi_{\boldsymbol{\theta}}(z_{p_i})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}. \tag{S.2}$$

Remark 1. *For the special case when $g = 0$, the closed form expressions of $U_n(\boldsymbol{\theta})$ and $I_n(\boldsymbol{\theta})$ can be obtained by taking the limit of $g \to 0$.*

Remark 2. *For a real data set $\{y_1, \ldots, y_n\}$, the corresponding z_{p_i}'s are unknown. We can plug their approximations $\tilde{z}_{p_i,k}$'s defined in (6) back into (S.1) and (S.2).*

2 Additional simulation results

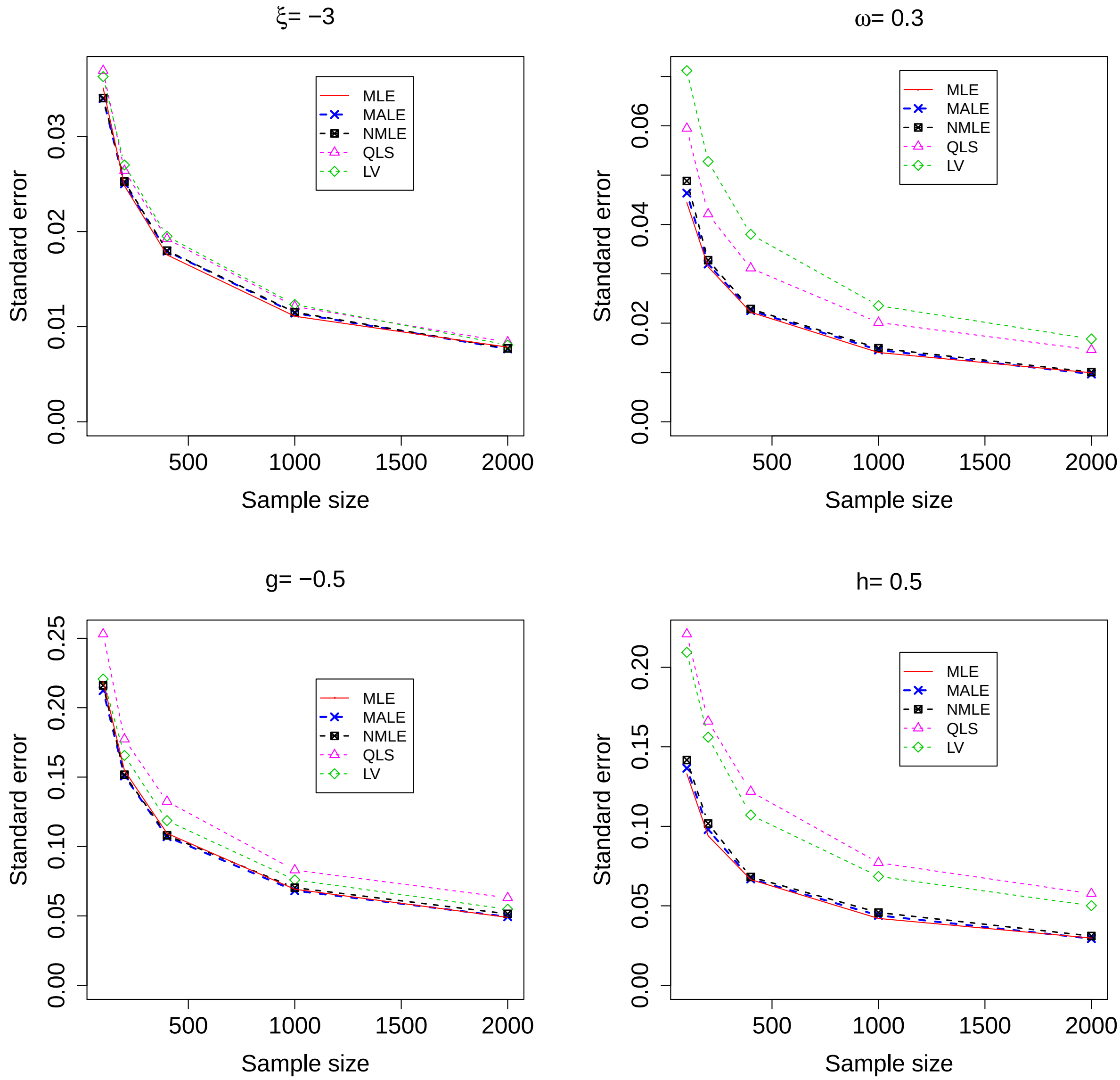

Figure 1: The estimation efficiencies of four methods for Tukey's g-and-h distribution; MLE: asymptotic standard error of the maximum likelihood estimator; MALE: maximum approximated likelihood estimator; NMLE: numerical maximum likelihood estimator; QLS: quantile least squares estimator; LV: letter-value-based estimator.

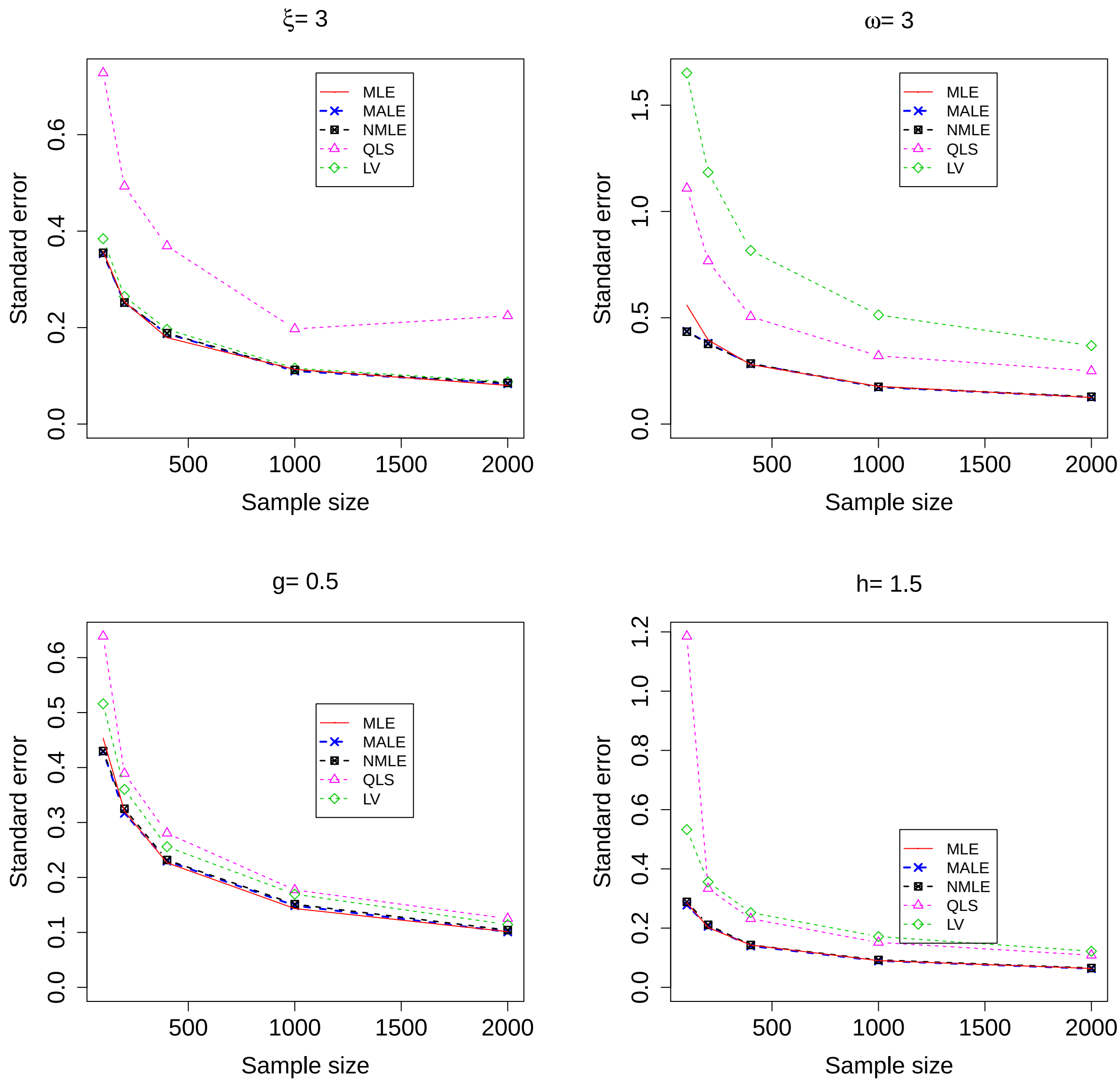

Figure 2: The estimation efficiencies of four methods for Tukey's g-and-h distribution; MLE: asymptotic standard error of the maximum likelihood estimator; MALE: maximum approximated likelihood estimator; NMLE: numerical maximum likelihood estimator; QLS: quantile least squares estimator; LV: letter-value-based estimator.

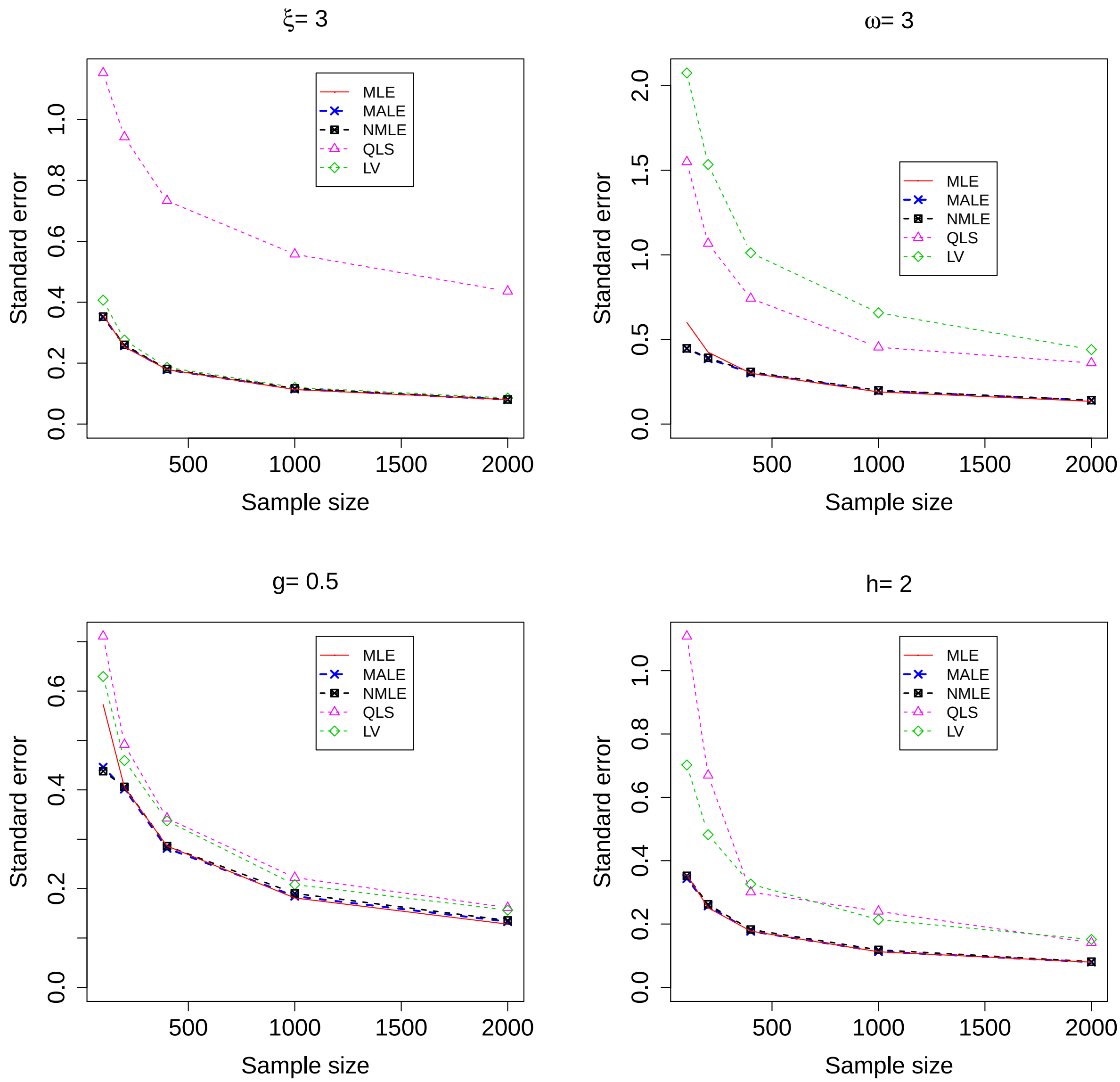

Figure 3: The estimation efficiencies of four methods for Tukey's g-and-h distribution; MLE: asymptotic standard error of the maximum likelihood estimator; MALE: maximum approximated likelihood estimator; NMLE: numerical maximum likelihood estimator; QLS: quantile least squares estimator; LV: letter-value-based estimator.

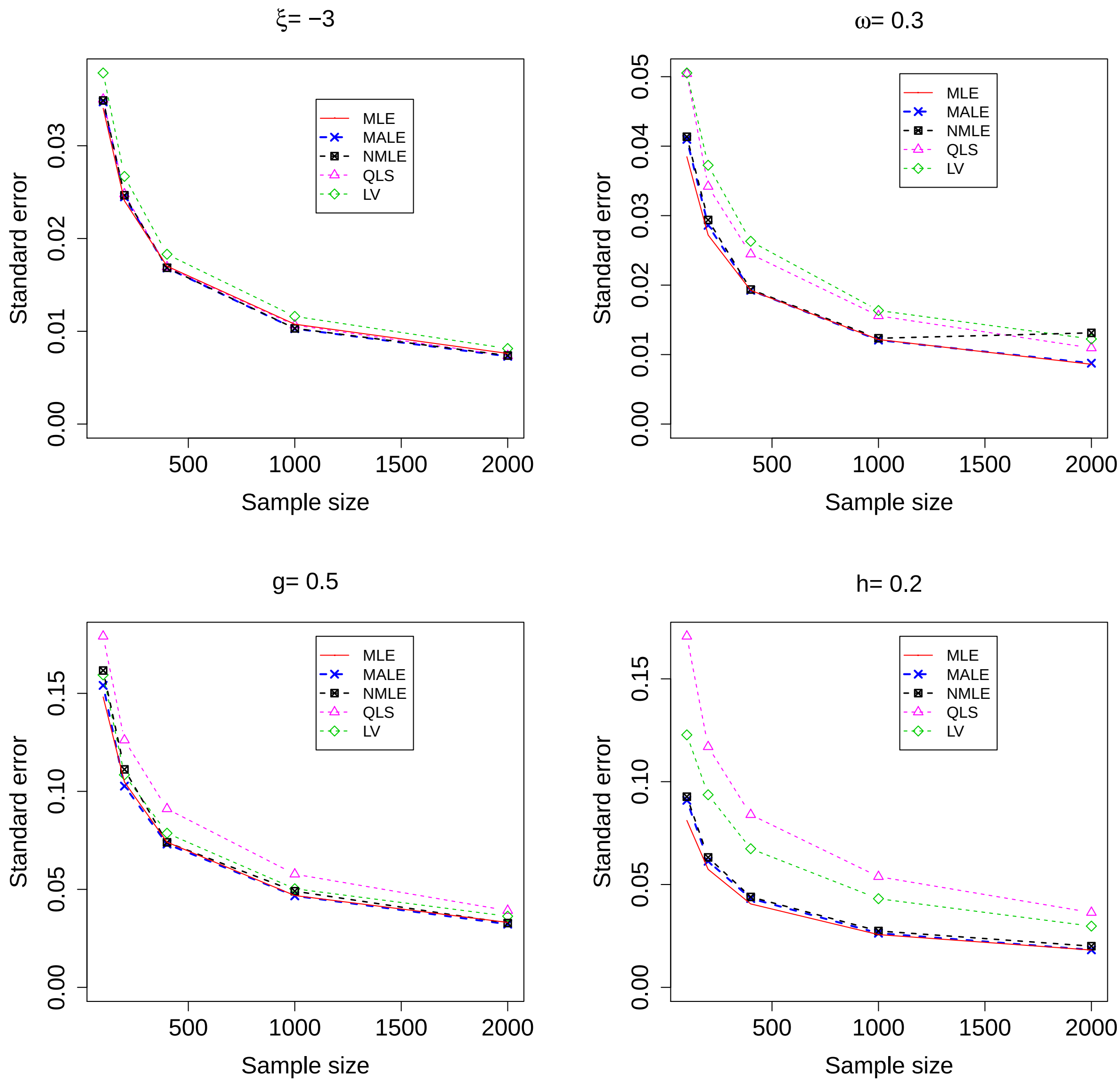

Figure 4: The estimation efficiencies of four methods for Tukey's g-and-h distribution; MLE: asymptotic standard error of the maximum likelihood estimator; MALE: maximum approximated likelihood estimator; NMLE: numerical maximum likelihood estimator; QLS: quantile least squares estimator; LV: letter-value-based estimator.

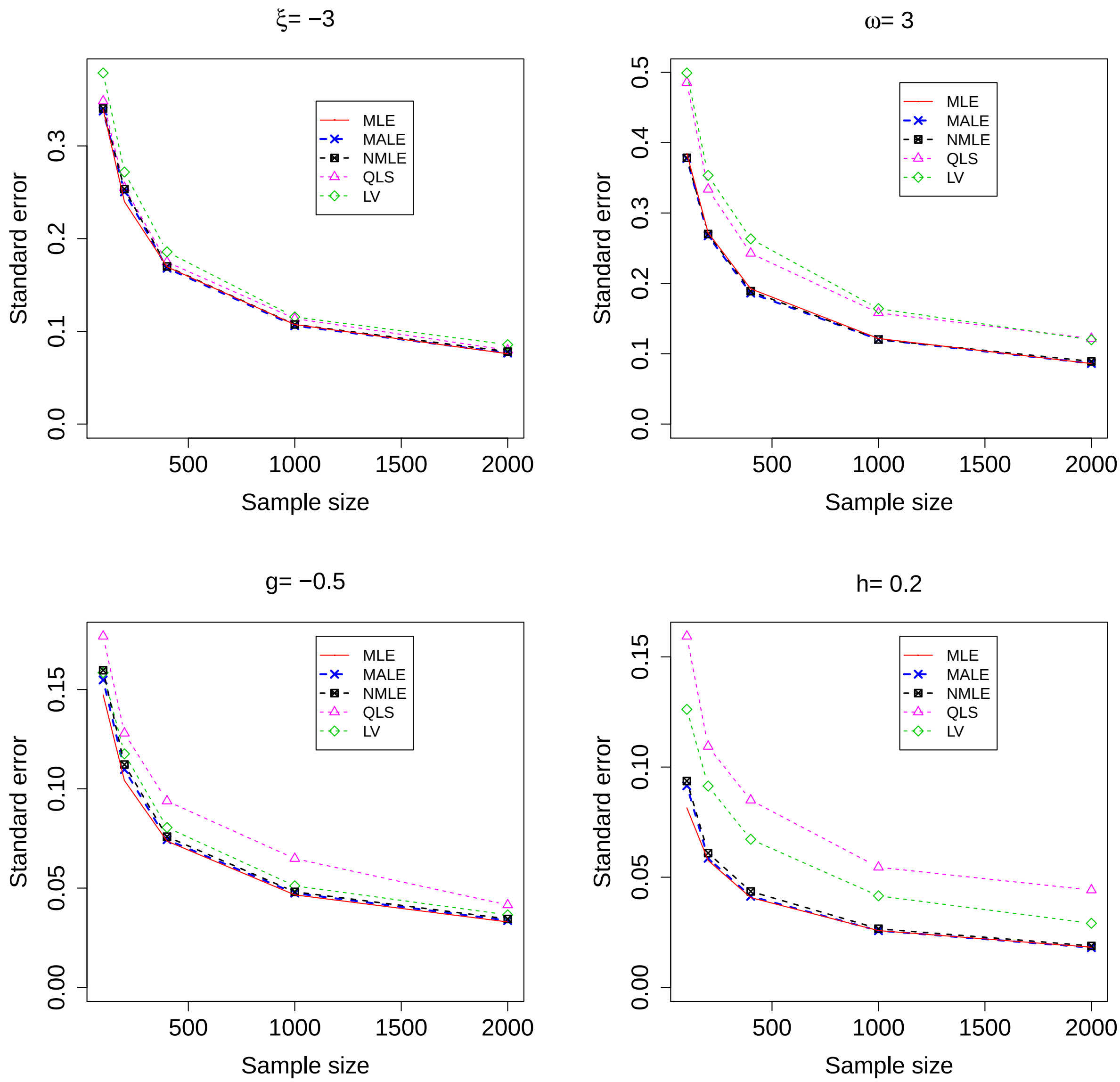

Figure 5: The estimation efficiencies of four methods for Tukey's g-and-h distribution; MLE: asymptotic standard error of the maximum likelihood estimator; MALE: maximum approximated likelihood estimator; NMLE: numerical maximum likelihood estimator; QLS: quantile least squares estimator; LV: letter-value-based estimator.

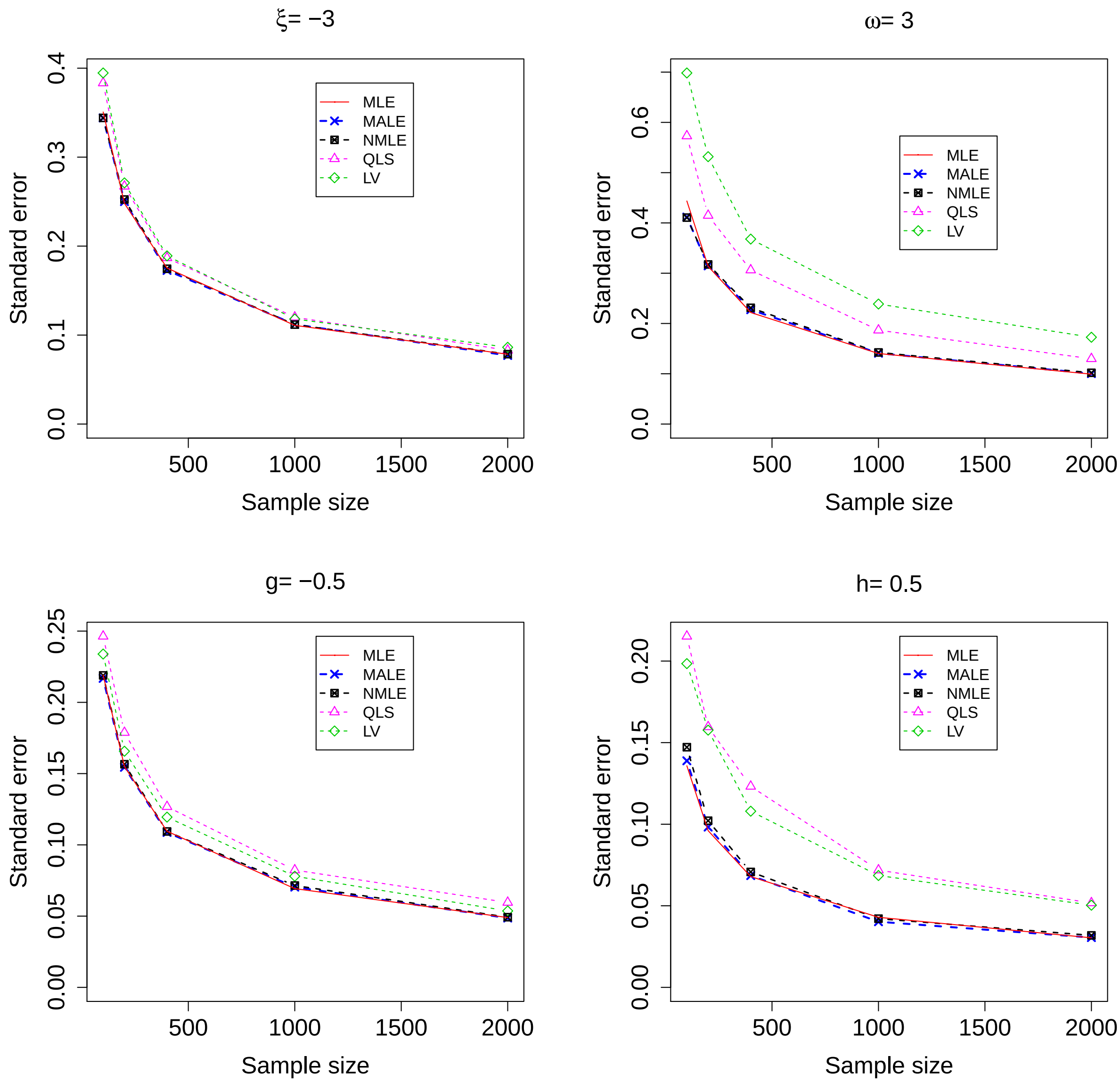

Figure 6: The estimation efficiencies of four methods for Tukey's g-and-h distribution; MLE: asymptotic standard error of the maximum likelihood estimator; MALE: maximum approximated likelihood estimator; NMLE: numerical maximum likelihood estimator; QLS: quantile least squares estimator; LV: letter-value-based estimator.