A Laplace Approximation for Approximate Bayesian Model Selection Richard M. Golden (**golden@utdallas.edu**), Shaurabh Nandy, Vishal Patel, and Pratibha Viraktamath

 $\mathbf{x}_1,...,\mathbf{x}_n$ be a realization of an *i.i.d.* sequence $\mathcal{D}_n \equiv [\tilde{\mathbf{x}}_1,...,\tilde{\mathbf{x}}_n]$ with common density $p_o(\mathbf{x})$ $\{p(\mathbf{x} | \boldsymbol{\theta}, \mathcal{M}) : \boldsymbol{\theta} \in \Theta_M \subset \mathcal{R}^q\}$ $(\mathcal{D}_{n} | \theta, \mathcal{M}) \equiv \prod p(\mathbf{x}_{i} | \theta, \mathcal{M})$ $(\boldsymbol{\theta};\mathcal{M})$ **Theory**
 $\overline{\mathbf{I}}$ $[\mathbf{x}_1, ..., \mathbf{x}_n]$ be a realization of an *i.i.d.* sequence $\tilde{\mathbf{D}}_n = [\tilde{\mathbf{x}}_1, ..., \tilde{\mathbf{x}}_n]$ with common density p_o 1 **Solution of an i.i.d. sequence** $\tilde{\mathbf{p}}_n = [\tilde{\mathbf{x}}_1, ..., \tilde{\mathbf{x}}_n]$ **be a realization of an** *i.i.d.* **sequence** $\tilde{\mathbf{p}}_n = [\tilde{\mathbf{x}}_1, ..., \tilde{\mathbf{x}}_n]$ **with common density** $p_n(\mathbf{x})$ **.
Probability Model : \mathbf{M} = \{p(\mathbf{x} | \mathbf{** $\left[\mathbf{x}_n\right]$ be a realization
 $\mathbf{M} \equiv \left\{ p(\mathbf{x} | \mathbf{\theta}, \mathbf{M}) : \mathbf{M} \right\}$ $[\mathbf{x}_n]$ be a realization of an *i.i.d.* sequently $\mathcal{M} = \{ p(\mathbf{x} | \mathbf{\theta}, \mathcal{M}) : \mathbf{\theta} \in \Theta_M \subset \mathbb{R}^q \}$
for $\mathcal{M}: p(\mathcal{D}_n | \mathbf{\theta}, \mathcal{M}) = \prod_{i=1}^n p(\mathbf{x}_i | \mathbf{\theta}, \mathcal{M})$ bability Model

elihood Functio
 $; M) \equiv -(1/n)1$ $p(\mathbf{x} | \boldsymbol{\theta}, \mathcal{M})$: $\boldsymbol{\theta} \in \Theta_{\mathcal{M}} \subset \mathcal{R}^q$ *n* $(\theta, \mathcal{M}) : \theta \in \Theta_M \subset \mathcal{F}$
 $(\theta, \mathcal{M}) \equiv \prod_{i=1}^n p(\mathbf{x}_i)$ *n* a realization of an *i.i.*
 $(\mathbf{x} | \mathbf{\theta}, \mathcal{M}) : \mathbf{\theta} \in \Theta_{\mathcal{M}}$ c
 $p(\mathcal{D}_n | \mathbf{\theta}, \mathcal{M}) \equiv \prod_{i=1}^n p_i$ **Probability Model :** $M = \{p(\mathbf{x})\}$
Likelihood Function for M : $p(\mathbf{x}) = \tilde{I}_n(\mathbf{0}; M) = -(1/n) \log p(\mathcal{D}_n | \mathbf{0}, \mathbf{0})$ $=$ **Probability Model :** $M = \{p(\mathbf{x} | \theta, M) : \theta \in \Theta_M \subset \mathbb{R}^q\}$
 Probability Model : $M = \{p(\mathbf{x} | \theta, M) : \theta \in \Theta_M \subset \mathbb{R}^q\}$ $\equiv \prod$ **Likelihood Function** for $M: p(\mathbf{D}, \mathbf{H}) = \prod_{i=1}^{n} p(\mathbf{x} | \mathbf{B}, \mathbf{M})$
 Likelihood Function for $M: p(\mathbf{D}_n | \mathbf{\Theta}, \mathbf{M}) = \prod_{i=1}^{n} p(\mathbf{x}_i | \mathbf{\Theta}, \mathbf{M})$ θ ; \mathcal{M}) = -(1/n) log p(\mathcal{D}_n | θ , \mathcal{M}), $\dot{\theta}_n$ = arg min $l_n(\theta; \mathcal{M})$ **x**_n] be a realization of an *i.i.d.* s
 $M = \{p(\mathbf{x} | \mathbf{\theta}, M) : \mathbf{\theta} \in \Theta_M \subset \mathcal{R}\}$ $(\theta; \mathcal{M}) \equiv E\left\{ \tilde{l}_n(\theta; \mathcal{M}) \right\}, \, \theta^* \equiv \arg \min l\left(\theta; \mathcal{M}\right)$ $(\theta_n;\mathcal{M}), \mathbf{B}_n \equiv (1/n)\sum \nabla \log p(\mathbf{x}_i | \theta_n, \mathcal{M})(\nabla \log p(\mathbf{x}_i | \theta_n, \mathcal{M}))$ $\big(\boldsymbol{\mathcal{D}}_{n} \mid \boldsymbol{\theta},\boldsymbol{\mathcal{M}}\big)=\exp\big(\!-\!nl_{n}\big(\boldsymbol{\theta};\boldsymbol{\mathcal{M}}\big)\big),\;\;p\big(\boldsymbol{\mathcal{D}}_{\!n}\mid\boldsymbol{\theta},\boldsymbol{\mathcal{M}}\big)=\exp\big(\!-\!nl\big(\boldsymbol{\theta};\boldsymbol{\mathcal{M}}\big)\big)$ $(\mathcal{D}_n | \mathcal{M}) \equiv \int p(\mathcal{D}_n | \boldsymbol{\theta}, \mathcal{M}) p_{\theta}(\boldsymbol{\theta} | \mathcal{M})$ 2 1 $\hat{\bm{\theta}}$: $\mathcal{M} = \{ p(\mathbf{x} | \boldsymbol{\theta}, \mathcal{M}) : \boldsymbol{\theta} \in \Theta_{\mathcal{M}} \subset \mathcal{R}^d \}$

on for $\mathcal{M}: p(\mathcal{D}_n | \boldsymbol{\theta}, \mathcal{M}) = \prod_{i=1}^n p(\mathbf{x}_i | \text{log } p(\mathcal{D}_n | \boldsymbol{\theta}, \mathcal{M}), \hat{\boldsymbol{\theta}}_n \equiv \argmin_{\boldsymbol{\theta} \in \Theta_{\mathcal{M}}} \tilde{l}_n(\boldsymbol{\theta};$ celihood Function for *M*: $p(\mathcal{D}_n | \theta, \mathcal{M})$
 $(\theta; \mathcal{M}) \equiv -(1/n) \log p(\mathcal{D}_n | \theta, \mathcal{M}), \hat{\theta}_n \equiv$
 $(\mathcal{M}) \equiv E\{\tilde{l}_n(\theta; \mathcal{M})\}, \theta^* \equiv \argmin_{\theta \in \Theta_{\mathcal{M}}} l(\theta; \theta)$ $-(1/n)\log p(\mathcal{D}_n | \theta, \mathcal{M}), \hat{\theta}_n = \argmin_{\theta \in \Theta_{\mathcal{M}}} \tilde{l}_n(\theta; \mathcal{M})$
 $\{\tilde{l}_n(\theta; \mathcal{M})\}, \theta^* = \argmin_{\theta \in \Theta_{\mathcal{M}}} l(\theta; \mathcal{M})$
 $\hat{\theta}_n; \mathcal{M}, \tilde{B}_n = (1/n)\sum_{i=1}^n \nabla \log p(\mathbf{x}_i | \hat{\theta}_n, \mathcal{M}) (\nabla \log p(\mathbf{x}_i | \hat{\theta}_n, \mathcal{M}))^T$ $\begin{aligned} \mathcal{U} & \equiv E\left\{\tilde{l}_n\left(\boldsymbol{\theta};\mathcal{M}\right)\right\}, \ \boldsymbol{\theta}^* \equiv \argmin_{\boldsymbol{\theta} \in \Theta_{\mathcal{M}}} l\left(\boldsymbol{\theta};\mathcal{M}\right) \ \nabla^2 \tilde{l}_n\left(\hat{\boldsymbol{\theta}}_n;\mathcal{M}\right), \ \tilde{\mathbf{B}}_n & \equiv (1/n)\sum_{i=1}^n \nabla \log p\left(\mathbf{x}_i \mid \hat{\boldsymbol{\theta}}_n, \mathcal{M}\right)\right)\nabla \log p \left(\mathbf{x}_i \mid \hat{\boldsymbol{\theta}}_n, \mathcal$ $f_n = (1/n) \sum_{i=1}^n \nabla \log p(\mathbf{x}_i | \hat{\theta}_n, \mathcal{M}) (\nabla \log p(\mathbf{x}_n))$
 $f_n(\theta; \mathcal{M})$, $p(\mathcal{D}_n | \theta, \mathcal{M}) = \exp(-nl(\theta; \mathcal{M}))$

for $\mathcal{M}: p(\mathcal{D}_n | \mathcal{M}) \equiv \int p(\mathcal{D}_n | \theta, \mathcal{M}) p_{\theta}(\theta)$ $M: p(\mathcal{D}_n | \theta, \mathcal{M}) \equiv \prod_{i=1}^n p(\theta_i | \theta, \mathcal{M}), \ \hat{\theta}_n \equiv \argmin_{\theta \in \Theta_{\mathcal{M}}} \tilde{l}_n$ *n n T* $\begin{aligned} \mathbf{\Theta};\mathbf{\mathcal{M}})\equiv E\left\{l_n\left(\mathbf{\Theta};\mathbf{\mathcal{M}}\right)\right\},\,\mathbf{\Theta}^*\equiv\argmin_{\mathbf{\Theta}\in\Theta_{\mathbf{\mathcal{M}}}}l\left(\mathbf{\Theta};\mathbf{\mathcal{M}}\right)\,,\\ m\equiv\nabla^2\tilde{l}_n\left(\hat{\mathbf{\Theta}}_n;\mathbf{\mathcal{M}}\right),\,\tilde{\mathbf{B}}_n\equiv(1/n)\sum_{i=1}^n\nabla\log p\left(\mathbf{x}_i\mid\hat{\mathbf{\Theta}}_n,\mathbf{\mathcal{M}}\right)\right)\left(\nabla\log$ *i* $\begin{aligned} &\mathbf{F} \nabla^2 \tilde{l}_n \left(\hat{\boldsymbol{\theta}}_n; \boldsymbol{\mathcal{M}} \right), \ \tilde{\mathbf{B}}_n \equiv (1/n) \sum_{i=1}^n \nabla \log p_i, \ &\mathbf{B}_n \mid \boldsymbol{\theta}, \boldsymbol{\mathcal{M}} \big) = \exp \left(-n \tilde{l}_n \left(\boldsymbol{\theta}; \boldsymbol{\mathcal{M}} \right) \right), \ \ p \left(\tilde{\boldsymbol{\mathcal{D}}}_n \right). \end{aligned}$ $p\left(\ddot{\mathbf{\mathcal{D}}}_n \mid \mathbf{\theta}, \mathbf{\mathcal{M}}\right) =$
 $\mathbf{p}_n \mid \mathbf{\mathcal{M}}$ = $\int p\left(\mathbf{\mathcal{D}}_n\right)$ $M \equiv \Big\{ p(\mathbf{x} | \boldsymbol{\theta}, \mathcal{M}) : \boldsymbol{\theta} \in \Theta_M \subset \text{for } \mathcal{M}: p(\mathcal{D}_n | \boldsymbol{\theta}, \mathcal{M}) \equiv \prod_{i=1}^n p_i \right\}$
 $p(\mathcal{D}_n | \boldsymbol{\theta}, \mathcal{M}), \hat{\boldsymbol{\theta}}_n \equiv \argmin_{\boldsymbol{\theta} \in \Theta_M} \hat{l}$ **Likelihood Function** for *M*: $p(\mathcal{D}_n | \theta)$
 $\tilde{l}_n(\theta; \mathcal{M}) \equiv -(1/n) \log p(\mathcal{D}_n | \theta, \mathcal{M}), \hat{\theta}$
 $l(\theta; \mathcal{M}) \equiv E\{\tilde{l}_n(\theta; \mathcal{M})\}, \theta^* \equiv \argmin_{\theta \in \Theta_{\mathcal{M}}} l$ $\hat{\theta}_n \equiv -(1/n) \log p(\mathcal{D}_n | \theta, \mathcal{M}), \ \hat{\theta}_n \equiv \argmin_{\theta \in \Theta_{\mathcal{M}}} \tilde{l}_n(\theta; \mathcal{M})$
 $\equiv E\left\{\tilde{l}_n(\theta; \mathcal{M})\right\}, \ \hat{\theta}^* \equiv \argmin_{\theta \in \Theta_{\mathcal{M}}} l(\theta; \mathcal{M})$
 $\tilde{l}_n(\hat{\theta}_n; \mathcal{M}), \ \tilde{\mathbf{B}}_n \equiv (1/n) \sum_{i=1}^n \nabla \log p(\mathbf{x}_i | \hat{\theta}_n, \mathcal{M}) (\nabla$ $\begin{aligned} &\mathcal{P}_n(\mathbf{\theta};\mathcal{M})\equiv E\left\{\tilde{l}_n\left(\mathbf{\theta};\mathcal{M}\right)\right\}, \ \mathbf{\theta}^* \equiv \argmin_{\mathbf{\theta}\in\Theta_{\mathcal{M}}} l\left(\mathbf{\theta};\mathcal{M}\right) \ &\tilde{\mathbf{A}}_n\equiv\nabla^2 \tilde{l}_n\left(\hat{\mathbf{\theta}}_n;\mathcal{M}\right), \ \tilde{\mathbf{B}}_n\equiv(1/n)\sum_{i=1}^n\nabla\log p\left(\mathbf{x}_i\mid\hat{\mathbf{\theta}}_n,\mathcal{M}\right)\right|\nabla\log p \ &$ $p(\mathbf{x}_i \mid \hat{\boldsymbol{\theta}}_n, \mathcal{M}) \Big(\nabla \log \mathbf{p} \Big(\mathbf{x}_i \mid \hat{\boldsymbol{\theta}}_n, \mathcal{M} \Big) \Big(\nabla \log \mathbf{p} \Big)$
 d)), $p(\mathcal{D}_n \mid \boldsymbol{\theta}, \mathcal{M}) = \exp(-nl)$
 $p(\mathcal{D}_n \mid \mathcal{M}) \equiv \int p(\mathcal{D}_n \mid \boldsymbol{\theta}, \mathcal{M}) p_{\theta}$ $\in \Theta$ **θ** $=$ \equiv **ood Function** for \mathcal{M} : $p(\mathcal{D}_n | \theta, \mathcal{M}) \equiv \prod_{i=1}^n$
 $\theta = -(1/n) \log p(\mathcal{D}_n | \theta, \mathcal{M}), \hat{\theta}_n \equiv \argmin_{\theta \in \Theta_{\mathcal{M}}}$
 $\theta \equiv E\left\{\tilde{l}_n(\theta; \mathcal{M})\right\}, \theta^* \equiv \argmin_{\theta \in \Theta_{\mathcal{M}}} l(\theta; \mathcal{M})$ $\begin{split} \mathbf{D}; \mathcal{M}) &\equiv -(1/n) \log p\left(\mathbf{\mathcal{D}}_n \mid \boldsymbol{\theta}, \mathcal{M}\right), \ \hat{\boldsymbol{\theta}}_n \equiv \argmin_{\boldsymbol{\theta} \in \Theta_{\mathcal{M}}} \tilde{l}_n\left(\boldsymbol{\theta}; \mathcal{M}\right), \\ \mathcal{M}) &\equiv E\left\{\tilde{l}_n\left(\boldsymbol{\theta}; \mathcal{M}\right)\right\}, \ \boldsymbol{\theta}^* \equiv \argmin_{\boldsymbol{\theta} \in \Theta_{\mathcal{M}}} l\left(\boldsymbol{\theta}; \mathcal{M}\right), \\ &\equiv \nabla^2 \tilde{l}_n\left(\hat$ $\begin{aligned} &\left\{\tilde{l}_n(\theta;\mathcal{M})\right\},\,\theta^*=\underset{\theta\in\Theta_{\mathcal{M}}}{\arg\min} \,l\left(\theta;\mathcal{M}\right)\\ &\left(\mathcal{M}\right),\,\tilde{\mathbf{B}}_n\equiv\left(1/n\right)\!\sum_{i=1}^n\nabla\log p\left(\mathbf{x}_i\mid\hat{\theta}_n,\mathcal{M}\right)\!\left(\nabla\log p\left(\mathbf{x}_i\mid\hat{\theta}_n,\mathcal{M}\right)\right)^T\\ &=\exp\left(-n\tilde{l}_n\left(\theta;\mathcal{M}\right)\right),\;\;p\left(\tilde{\mathbf{B$ \sum **θ** $\begin{aligned} &\mathcal{P}(\mathbf{X} | \boldsymbol{\theta}, \boldsymbol{\mathcal{M}}) : \boldsymbol{\theta} \in \Theta_{\boldsymbol{\mathcal{M}}} \subset \mathcal{R}^q \right\} \ &\mathcal{P}(\mathcal{D}_n | \boldsymbol{\theta}, \boldsymbol{\mathcal{M}}) \equiv \prod_{i=1}^n p(\mathbf{x}_i | \boldsymbol{\theta}, \boldsymbol{\mathcal{M}}) \ &\mathbf{\theta}, \boldsymbol{\mathcal{M}}), \ \hat{\boldsymbol{\theta}}_n \equiv \argmin_{\boldsymbol{\theta} \in \Theta_{\mathcal{M}}} \tilde{l}_n(\boldsymbol{\theta}; \boldsymbol{\mathcal{M}}) \end{aligned}$ **ikelihood Function** for \mathcal{M} : $p(\mathcal{D}_n | \theta, \mathcal{M}) = \prod_{i=1}^n p(\mathbf{x}_i | \theta, \mathcal{M})$
 $(\theta; \mathcal{M}) = -(1/n) \log p(\mathcal{D}_n | \theta, \mathcal{M}), \ \hat{\theta}_n = \argmin_{\theta \in \Theta_{\mathcal{M}}} \tilde{l}_n(\theta; \mathcal{M})$
 $\theta; \mathcal{M}) \equiv E\{\tilde{l}_n(\theta; \mathcal{M})\}, \ \theta^* \equiv \argmin_{\theta \in \Theta_{\mathcal{M}}} l(\theta;$ $\tilde{l}_n(\theta; \mathcal{M}) = -(1/n) \log p(\mathcal{D}_n | \theta, \mathcal{M}), \ \hat{\theta}_n \equiv \argmin_{\theta \in \Theta_{\mathcal{M}}} \tilde{l}_n(\theta; \mathcal{M})$
 $l(\theta; \mathcal{M}) \equiv E\{\tilde{l}_n(\theta; \mathcal{M})\}, \ \theta^* \equiv \argmin_{\theta \in \Theta_{\mathcal{M}}} l(\theta; \mathcal{M})$
 $\tilde{\mathbf{A}}_n \equiv \nabla^2 \tilde{l}_n(\hat{\theta}_n; \mathcal{M}), \ \tilde{\mathbf{B}}_n \equiv (1/n) \sum_{i=1}^n \$ $\begin{aligned} \mathbf{H} &\equiv \mathbf{E} \left\{ \tilde{l}_n \left(\mathbf{\theta}; \mathbf{\mathcal{M}} \right) \right\}, \, \mathbf{\theta}^* \equiv \argmin_{\mathbf{\theta} \in \Theta_{\mathcal{M}}} l \left(\mathbf{\theta}; \mathbf{\mathcal{M}} \right) \end{aligned}$
 $\mathbf{H}^2 \tilde{l}_n \left(\hat{\mathbf{\theta}}_n; \mathbf{\mathcal{M}} \right), \, \tilde{\mathbf{B}}_n \equiv (1/n) \sum_{i=1}^n \nabla \log p \left(\mathbf{x}_i \mid \hat{\mathbf{\theta}}_n, \mathbf{\$ $\tilde{\mathbf{A}}_n = \nabla^2 \tilde{l}_n \left(\hat{\mathbf{\theta}}_n; \mathcal{M} \right), \ \tilde{\mathbf{B}}_n = (1/n) \sum_{i=1}^n \nabla \log p \left(\mathbf{x}_i \mid \hat{\mathbf{\theta}}_n, \mathcal{M} \right) \left(\nabla \log p \left(\mathbf{x}_i \mid \hat{\mathbf{\theta}}_n, \mathcal{M} \right) \right)^T$
 $p(\mathcal{D}_n \mid \mathbf{\theta}, \mathcal{M}) = \exp \left(-n \tilde{l}_n \left(\mathbf{\theta}; \mathcal{M} \right) \right), \ p(\$ **M M** $\begin{aligned} \mathbf{E} & \Big\{ p\big(\mathbf{x} \mid \boldsymbol{\theta}, \mathcal{M}\big) : \boldsymbol{\theta} \in \Theta_{\mathcal{M}} \subset \mathcal{R}^q \Big\} \\ \mathcal{M}: & p\big(\mathcal{D}_n \mid \boldsymbol{\theta}, \mathcal{M}\big) \equiv \prod_{i=1}^n p\big(\mathbf{x}_i \mid \boldsymbol{\theta}, \mathcal{M}\big), \\ \mathcal{D}_n \mid \boldsymbol{\theta}, \mathcal{M}\big), & \hat{\boldsymbol{\theta}}_n \equiv \underset{\boldsymbol{\theta} \in \Theta_{\mathcal{M}}}{\arg \min} \tilde{l}_n\big(\boldsymbol{\theta}; \$ elihood Function for *M*: $p(\mathcal{D}_n | \theta, \mathcal{M}) =$
 $(\mathcal{M}) = -(1/n) \log p(\mathcal{D}_n | \theta, \mathcal{M}), \hat{\theta}_n = \text{ar}$
 $(\mathcal{M}) = E\{\tilde{l}_n(\theta; \mathcal{M})\}, \theta^* = \argmin_{\theta \in \Theta_{\mathcal{M}}} l(\theta; \mathcal{M})$ *I* n) $\log p(\mathcal{D}_n | \theta, \mathcal{M}), \hat{\theta}_n \equiv \argmin_{\theta \in \Theta_{\mathcal{M}}} \tilde{l}_n(\theta; \mathcal{M})$
 $(\theta; \mathcal{M})\}, \theta^* \equiv \argmin_{\theta \in \Theta_{\mathcal{M}}} l(\theta; \mathcal{M})$
 $\mathcal{M}, \tilde{\mathbf{B}}_n \equiv (1/n) \sum_{i=1}^n \nabla \log p(\mathbf{x}_i | \hat{\theta}_n, \mathcal{M}) (\nabla \log p(\mathbf{x}_i | \hat{\theta}_n, \mathcal{M}))$ $\begin{aligned} \mathbf{D}; \mathcal{M}) &= E\left\{\tilde{l}_n\left(\mathbf{\theta};\mathcal{M}\right)\right\}, \, \mathbf{\theta}^* \equiv \argmin_{\mathbf{\theta} \in \Theta_{\mathcal{M}}} l\left(\mathbf{\theta};\mathcal{M}\right) \ &= \nabla^2 \tilde{l}_n\left(\hat{\mathbf{\theta}}_n;\mathcal{M}\right), \, \tilde{\mathbf{B}}_n \equiv (1/n) \sum_{i=1}^n \nabla \log p\left(\mathbf{x}_i \mid \hat{\mathbf{\theta}}_n, \mathcal{M}\right) \left(\nabla \log p\left(\mathbf{x}_i\right$ 1/n) $\sum_{i=1}^{n} \nabla \log p(\mathbf{x}_i | \hat{\theta}_n, \mathcal{M}) (\nabla \log p(\mathbf{x}_i | \hat{\theta}_n, \mathcal{M}))^T$
 (i), $p(\tilde{\mathbf{Z}}_n | \theta, \mathcal{M}) = \exp(-nl(\theta; \mathcal{M}))$
 M: $p(\mathbf{Z}_n | \mathcal{M}) \equiv \int p(\mathbf{Z}_n | \theta, \mathcal{M}) p_{\theta}(\theta | \mathcal{M}) d\theta$ with prior $p_{\theta}(\theta | \mathcal{M})$
 $2\pi E [\log n$ $\left\{\hat{\theta}_n;\mathcal{M}\right\}+2q=-2nE\left\{\log p\left(\tilde{\mathbf{D}}_n\mid\boldsymbol{\theta}^*,\mathcal{M}\right)\right\}+o_p\left(1\right)$ $\left(\mathbf{\Theta}_n;\mathcal{M}\right) + 2TRACE \left(\left(\mathbf{A}_n\right) \quad \mathbf{B}_n\right) = -2nE \left\{ \log p\left(\mathbf{\mathcal{D}}_n \mid \mathbf{\Theta}^*, \mathbf{\mathcal{M}}\right) \right\} + o_p\left(1\right) \text{ (Takeuchi, 1976)}$ $(\theta_n;\mathcal{M}) + q \log(n) = -2 \log p(\mathcal{D}_n | \mathcal{M}) + O_p(1)$ (Schwarz, 1978) $\hat{\mathbf{\Theta}}$: M) + 2TRACE($(\tilde{\mathbf{A}})^{-1}$ * $\begin{pmatrix} y \\ y \end{pmatrix}$ with prior $p_{\theta}(\theta)$ $\begin{aligned} \n\mathcal{L} & \mathcal{L}_{\ell_n}(\sigma_n, \text{cyc}), \mathbf{D}_n = (\mathbf{I} \wedge \text{D}) \sum_{i=1}^{n} \nabla \log P(\mathbf{x}_i \mid \sigma_n, \text{cyc}) (\nabla \log P(\mathbf{x}_i \mid \sigma_n, \text{cyc})) \n\end{aligned}$ **10.** \mathcal{M}) = exp($-n\tilde{l}_n(\theta; \mathcal{M})$), $p(\tilde{\mathcal{B}}_n \mid \theta, \mathcal{M}) = \exp(-nl(\theta; \mathcal{M}))$
 ginal Li 1 Likelihood for $\mathcal{M}: p(\mathcal{D}_n | \mathcal{M}) = \int p(\mathcal{D}_n | \theta, \mathcal{M}) p_o(\theta | \mathcal{M}) d\theta$ with prior $p_o(\theta | \mathcal{M})$
 $\tilde{u}_n(\hat{\theta}_n; \mathcal{M}) + 2q = -2nE \{ \log p(\tilde{\mathcal{D}}_n | \theta^*, \mathcal{M}) \} + o_p(1)$ only if $p_o \in \mathcal{M}$ (Akaike, 1974)
 $2n\tilde{u}_n(\hat{\$ $\begin{aligned} \mathbf{E} &= 2n\tilde{l}_n\left(\hat{\mathbf{\theta}}_n;\mathcal{M}\right) + 2q = -2nE\left\{\log p\left(\tilde{\mathbf{\mathcal{D}}}_n \mid \mathbf{\theta}^*,\mathcal{M}\right)\right\} + o_p\left(1\right) \text{ only if } P_o \in \mathbb{C} \\ \mathbf{E} &= 2n\tilde{l}_n\left(\hat{\mathbf{\theta}}_n;\mathcal{M}\right) + 2TRACE\left(\left(\tilde{\mathbf{A}}_n\right)^{-1}\tilde{\mathbf{B}}_n\right) = -2nE\left\{\log p\left(\tilde{\mathbf$ $=\overline{2}$
 $=\overline{2}$ **Likelihood** for $\mathcal{M}: p(\mathcal{D}_n | \mathcal{M}) \equiv \int p(\mathcal{D}_n | \theta, \mathcal{M}) p_{\theta}(\theta | \mathcal{M}) d\theta$
 $\tilde{p}_n(\hat{\theta}_n; \mathcal{M}) + 2q = -2nE \{ \log p(\tilde{\mathcal{D}}_n | \theta^*, \mathcal{M}) \} + o_p(1)$ only if p_o $\left\{\hat{\mathbf{\Theta}}_n; \mathbf{\mathcal{M}}\right\} + 2q = -2nE\left\{\log p\left(\tilde{\mathbf{\mathcal{D}}}_n \mid \mathbf{\theta}^*, \mathbf{\mathcal{M}}\right)\right\} + o_p\left(1\right) \text{ only if } p_o \in \mathbf{\mathcal{M}}$
 $\tilde{\mathbf{\Theta}}_n; \hat{\mathbf{\mathcal{M}}}\right\} + 2TRACE\left(\left(\tilde{\mathbf{A}}_n\right)^{-1}\tilde{\mathbf{B}}_n\right) = -2nE\left\{\log p\left(\tilde{\mathbf{\mathcal{D}}}_n \mid \mathbf{\theta}^*, \math$ $2n\tilde{l}_n(\hat{\theta}_n;\mathcal{M}) + 2TRACE((\tilde{A}_n)^{-1}\tilde{B}_n) = -2nE$
 $\tilde{h}_n(\hat{\theta}_n;\mathcal{M}) + q \log(n) = -2\log p(\mathcal{D}_n | \mathcal{M}) + O_p$ $GBLC_{L} = 2nl_{n}(\theta_{n};M) - 2\log p_{\theta}(\theta_{n} | M) + q\log \left| \frac{n}{2}\right| + \log \det A_{n} = -2\log p(\mathcal{D}_{n} | M) + o_{p}(1)$ (\mathbf{d}, \mathbf{d}) *d* **b** with prior *p* **Likelihood** for $\mathcal{M}: p(\mathcal{D}_n | \mathcal{M}) = \int p(\mathcal{D}_n | \theta, \mathcal{M}) p_{\theta}(\theta | \mathcal{M}) d\theta$ with
 $\tilde{h}_n(\hat{\theta}_n; \mathcal{M}) + 2q = -2nE \{ \log p(\tilde{\mathcal{D}}_n | \theta^*, \mathcal{M}) \} + o_p(1)$ only if $p_{\theta} \in \mathcal{M}$
 $n\tilde{h}_n(\hat{\theta}_n; \mathcal{M}) + 2TRACE \left((\tilde{\mathbf{A}}_n)^{-1$ $2n\tilde{l}_n(\hat{\theta}_n;\mathcal{M}) + 2q = -2nE\left\{\log p\left(\tilde{\mathbf{D}_n} \mid \hat{\theta}^*, \mathcal{M}\right)\right\}$
= $2n\tilde{l}_n(\hat{\theta}_n;\mathcal{M}) + 2TRACE\left(\left(\tilde{\mathbf{A}}_n\right)^{-1}\tilde{\mathbf{B}}_n\right) = -2nE$
 $n\tilde{l}_n(\hat{\theta}_n;\mathcal{M}) + q\log(n) = -2\log p(\mathbf{D}_n \mid \mathcal{M}) + O$ θ **Marginal Likelihood** for $\mathcal{M}: p(\mathcal{D}_n | \mathbf{0}, \mathcal{M}) = \text{Exp}(-\mathcal{M}(\mathbf{0}, \mathcal{M}))$
 Marginal Likelihood for $\mathcal{M}: p(\mathcal{D}_n | \mathcal{M}) \equiv \int p(\mathcal{D}_n | \mathbf{\theta}, \mathcal{M}) p_{\theta}(\mathbf{\theta} | \mathcal{M}) d\mathbf{\theta}$ with prior $p_{\theta}(\mathbf{\theta} | \mathcal{M})$
 AIC $(\mathbf{\mathcal{M}})^T$
 θ with prior $p_{\theta}(\mathbf{\theta} | \mathbf{\mathcal{M}})$ **AIC** = $2n\tilde{l}_n(\hat{\theta}_n; \mathcal{M}) + 2q = -2nE\{log P(\hat{\theta}_n | \theta, \mathcal{M})\} + o_p(1)$ only
 AIC = $2n\tilde{l}_n(\hat{\theta}_n; \mathcal{M}) + 2q = -2nE\{log P(\hat{\theta}_n | \theta, \mathcal{M})\} + o_p(1)$ only $\mathbf{BIC} = 2nl_n$ Θ η prior η (θ | M) $\exp(-n\tilde{l}_n(\theta; \mathcal{M})), p(\tilde{\mathcal{D}}_n | \theta, \mathcal{M}) = \exp(-nl(\theta; \mathcal{M}))$
 (index)
 M and to **m** and $\mathcal{M}: p(\mathcal{D}_n | \mathcal{M}) \equiv \int p(\mathcal{D}_n | \theta, \mathcal{M}) p_{\theta}(\theta | \mathcal{M}) d\theta$ with
 \mathcal{M} + 2q = -2nE {log p($\tilde{\mathcal{D}}_n | \theta^*, \mathcal{M}$ } + o_p(1 $(\mathbf{M}) + 2q = -2nE \{ \log p \left(\mathbf{\tilde{D}}_n \mid \mathbf{\theta}^* \right) \}$
 $(\mathbf{\hat{\theta}}_n; \mathbf{M}) + 2TRACE \left(\left(\mathbf{\tilde{A}}_n \right)^{-1} \mathbf{\tilde{B}}_n \right) =$
 $(\mathbf{\hat{M}}) + q \log(n) = -2 \log p (\mathbf{\mathcal{D}}_n \mid \mathbf{M})$ $\left(\mathbf{\theta}_n;\mathcal{M}\right) - 2\log p_\theta\left(\mathbf{\theta}_n \mid \mathcal{M}\right) + q\log\left(\frac{n}{2\pi}\right) - \log\det\left(\left(\mathbf{A}_n\right) \mid \mathbf{B}_n\right) = -2\log p\left(\mathbf{\mathcal{D}}_n \mid \mathcal{M}\right) + o_p\left(1\right)\left(\text{Lv and Liu, 2014}\right)$ $(\theta_n; \mathcal{M}) - 2 \log p_\theta (\theta_n | \mathcal{M})$ 1 $\hat{\theta}_n$; \mathcal{M} + 2TRACE $((\tilde{A}_n)^{-1} \tilde{B}_n) = -2nE \{ \log p(\tilde{\mathcal{D}}_n | \theta^*, \mathcal{M}) \} + o_p(1)$ (Takeuchi, 1976)
 \mathcal{M} + $q \log(n) = -2 \log p(\mathcal{D}_n | \mathcal{M}) + O_p(1)$ (Schwarz, 1978)
 $\hat{\theta}_n$; \mathcal{M}) - 2log $p_\theta(\hat{\theta}_n | \mathcal{M}) + q \log \$ 2 $2n\tilde{l}_n(\hat{\theta}_n;\mathcal{M}) + q \log(n) = -2 \log p(\mathcal{D}_n | \mathcal{M}) + O_p(1)$ (Schwarz, 1978)
 $L = 2n\tilde{l}_n(\hat{\theta}_n;\mathcal{M}) - 2 \log p_\theta(\hat{\theta}_n | \mathcal{M}) + q \log\left(\frac{n}{2\pi}\right) + \log \det \tilde{A}_n = -2 \log p(\mathcal{D}_n | \mathcal{M}) + o_p(1)$ (e.g., Wasserman, 2000)
 $= 2n\tilde{l}_n(\hat{\theta}_n;\mathcal{M})$ 2 $=2n\tilde{l}_n(\hat{\theta}_n;\mathcal{M})-2\log p_\theta(\hat{\theta}_n\mid\mathcal{M})+q\log\left(\frac{n}{2\pi}\right)$
 $2n\tilde{l}_n(\hat{\theta}_n;\mathcal{M})-2\log p_\theta(\hat{\theta}_n\mid\mathcal{M})+q\log\left(\frac{n}{2\pi}\right)-$
 $=2n\tilde{l}_n(\hat{\theta}_n;\mathcal{M})-2\log p_\theta(\hat{\theta}_n\mid\mathcal{M})+q\log\left(\frac{n}{2\pi}\right)$ $\hat{\mathbf{D}}_n$; \mathcal{M} + q log(n) = -2 log p($\mathcal{D}_n \mid \mathcal{M}$ + O_p (1) (Schwarz, 1978)
 $\sum_{n} (\hat{\mathbf{\theta}}_n; \mathcal{M}) - 2 \log p_\theta (\hat{\mathbf{\theta}}_n \mid \mathcal{M}) + q \log \left(\frac{n}{2\pi}\right) + \log \det \tilde{\mathbf{A}}_n = -2 \log p(\mathcal{D}_n \mid \mathcal{M}) + o_p$ $n_n(\theta_n; \mathcal{M}) - 2\log p_\theta(\theta_n | \mathcal{M}) + q \log \left(\frac{1}{2\pi}\right) + \log \det \mathbf{A}_n = -2\log p(\mathbf{\mathcal{D}}_n | \mathcal{M}) + o_p(1)$ (e.good $\int_{n_n}^{\infty} (\hat{\theta}_n; \mathcal{M}) - 2\log p_\theta(\hat{\theta}_n | \mathcal{M}) + q \log \left(\frac{n}{2\pi}\right) - \log \det \left((\tilde{\mathbf{A}}_n)^{-1} \tilde{\mathbf{B}}_n\right) = -2\log p(\mathbf{\mathcal{D}}_n | \mathcal{$ $=2nl_n(\theta_n;\mathcal{M})-2\log p_\theta(\theta_n)$
 $p_p=2n\tilde{l}_n(\hat{\theta}_n;\mathcal{M})-2\log p_\theta(\hat{\theta}_n)$ *n* $n\tilde{l}_n(\hat{\theta}_n;\mathcal{M}) + 2TRACE\left((\tilde{\mathbf{A}}_n)^{-1}\tilde{\mathbf{B}}_n\right) = -2nE\left\{\log p\left(\tilde{\mathbf{D}}_n \mid \theta^*, \mathcal{M}\right)\right\} + o_p(1)$ (Takeu
 $(\hat{\theta}_n;\mathcal{M}) + q\log(n) = -2\log p(\mathbf{D}_n \mid \mathcal{M}) + O_p(1)$ (Schwarz, 1978)
 $n\tilde{l}_n(\hat{\theta}_n;\mathcal{M}) - 2\log p_\theta(\hat{\theta}_n \mid \mathcal{M}) +$ *n* $\begin{aligned} &\nabla_n \left(\hat{\theta}_n; \mathcal{M} \right) + q \log(n) = -2 \log p \left(\mathcal{D}_n \mid \mathcal{M} \right) + O_p \left(1 \right) \text{ (Schwarz,1978)}\\ &2n\tilde{l}_n \left(\hat{\theta}_n; \mathcal{M} \right) - 2 \log p_\theta \left(\hat{\theta}_n \mid \mathcal{M} \right) + q \log \left(\frac{n}{2\pi} \right) + \log \det \tilde{A}_n = -2 \log p \left(\mathcal{D}_n \mid \mathcal{M} \right) + o_p \left(1 \right) \text{ (e)}\\ &n$ $n l_n \left(\mathbf{\theta}_n;\mathbf{\mathcal{M}}\right) - 2\log p_\theta \left(\mathbf{\theta}_n \mid \mathbf{\mathcal{M}}\right) + q$
 $\tilde{l}_n \left(\hat{\mathbf{\theta}}_n;\mathbf{\mathcal{M}}\right) - 2\log p_\theta \left(\hat{\mathbf{\theta}}_n \mid \mathbf{\mathcal{M}}\right) + q$ let θ θ θ (θ _n | θ) θ | θ | θ | θ | θ $\begin{aligned} \left\{ \log p\left(\mathbf{\tilde{D}}_{n} | \boldsymbol{\theta}^{*}, \mathbf{\mathcal{M}}\right) \right\} \\ \left\{ \log p\left(\mathbf{\tilde{D}}_{n} | \boldsymbol{\theta}^{*}, \mathbf{\mathcal{M}}\right) \right\} \\ \left\{ \log p\left(1\right) \left(\text{Schwarz}, 1978\right) \right\} \\ \left\{ \frac{n}{2\pi}\right\} + \log \det \tilde{\mathbf{A}}_{n} = -2\log n, \end{aligned}$ ${\cal T}$ —
— $\hat{\theta}_n$; M) + 2TRACE $((\tilde{A}_n)^{-1} \tilde{B}_n) = -2nE \{ \log p(\tilde{D}_n | \theta^*, M) \} + o_p(1)$ (Takeuchi, 1976)
 M) + q log(n) = -2 log p($\mathcal{D}_n | M$) + $O_p(1)$ (Schwarz, 1978)
 $\hat{\theta}_n$; M) - 2 log $p_\theta(\hat{\theta}_n | M) + q \log \left(\frac{n}{2\pi} \right)$ + \mathcal{U} + $O_p(1)$ (Schwarz, 1978)
g $\left(\frac{n}{2\pi}\right)$ + log det $\tilde{A}_n = -21$
 $\left(\frac{n}{2\pi}\right)$ - log det $\left(\tilde{A}_n\right)^{-1} \tilde{B}_n$ $q \log(n) = -2 \log p(\mathcal{D}_n | \mathcal{M}) + O_p(1)$ (Schwarz, 1978)
 $-2 \log p_\theta(\hat{\theta}_n | \mathcal{M}) + q \log \left(\frac{n}{2\pi}\right) + \log \det \tilde{A}_n = -2 \log p(\mathcal{D}_n | \mathcal{M}) + o_p(1)$ (e.g., Wasserman, 20
 $-2 \log p_\theta(\hat{\theta}_n | \mathcal{M}) + q \log \left(\frac{n}{2\pi}\right) - \log \det \left((\tilde{A}_n)^{-1} \tilde{B}_n\right) = -2 \log$ $-2\log p_{\theta}\left(\hat{\theta}_n \mid \mathcal{M}\right) + q \log \left(\frac{n}{2\pi}\right) + \log \det$
 $2\log p_{\theta}\left(\hat{\theta}_n \mid \mathcal{M}\right) + q \log \left(\frac{n}{2\pi}\right) - \log \det$
 $-2\log p_{\theta}\left(\hat{\theta}_n \mid \mathcal{M}\right) + q \log \left(\frac{n}{2\pi}\right) - \log \det$ **GBIC** $= 2n\tilde{l}_n(\hat{\theta}_n; \mathcal{M}) + q \log(n) = -2 \log p(\mathcal{D}_n | \mathcal{M}) + O_p(1)$ (Schwarz, 1978)
 GBIC $_L = 2n\tilde{l}_n(\hat{\theta}_n; \mathcal{M}) - 2 \log p_\theta(\hat{\theta}_n | \mathcal{M}) + q \log \left(\frac{n}{2\pi}\right) + \log \det \tilde{A}_n = -2 \log p(\mathcal{D}_n | \mathcal{M}) + o_p$
 GBIC $= 2n\tilde{l}_n(\hat{\theta}_n; \mathcal{M}) -$ **GBIC**_L=2nl_n $(\theta_n; \mathcal{M})$ - 2 log p_θ $(\theta_n | \mathcal{M})$ + q log $\left(\frac{n}{2\pi}\right)$
 GBIC=2nl_n $(\hat{\theta}_n; \mathcal{M})$ - 2 log p_θ $(\hat{\theta}_n | \mathcal{M})$ + q log $\left(\frac{n}{2\pi}\right)$ -
 GBIC_P=2nl_n $(\hat{\theta}_n; \mathcal{M})$ - 2 log p_θ $(\hat$ **t**) + $q \log(n) = -2 \log p(\mathcal{D}_n \mid \mathcal{M}) + O_p(1)$ (Schwarz, 1978)

; \mathcal{M}) - $2 \log p_\theta(\hat{\theta}_n \mid \mathcal{M}) + q \log\left(\frac{n}{2\pi}\right) + \log \det \tilde{A}_n = -2 \log p(\mathcal{D}_n \mid \mathcal{M}) + o_p(1)$
 \mathcal{M}) - $2 \log p_\theta(\hat{\theta}_n \mid \mathcal{M}) + q \log\left(\frac{n}{2\pi}\right) - \log \det \left((\tilde{A}_n)^{-1} \$ M $-2\log p_\theta(\theta_n | \mathcal{M}) + q \log \left(\frac{1}{2\pi} \right) + \log \det \mathbf{A}_n = -2\log p(\mathcal{D}_n | \mathcal{M}) + o_p(1)$ (e.
 M $-2\log p_\theta(\hat{\theta}_n | \mathcal{M}) + q \log \left(\frac{n}{2\pi} \right) - \log \det \left(\left(\tilde{\mathbf{A}}_n \right)^{-1} \tilde{\mathbf{B}}_n \right) = -2\log p(\mathcal{D}_n | \mathcal{M}) + o_p(1)$
 M $-2\log p_\theta(\hat{\theta}_n | \mathcal{$ $= -2 \log p(\mathcal{D}_n | \mathcal{M}) + o_p(1)$ (Lv and Liu, 2014) $\left(\theta_n;\mathcal{M}\right)-2\log p_\theta\left(\theta_n\mid\mathcal{M}\right)+q\log\left(\frac{N}{2\pi}\right)+\log\det A_n+TRACE\left(\left(\mathbf{A}_n\right)\right)\mathbf{B}_n\right)$ $= -2 \log p(\ddot{\mathbf{D}}_n | \mathbf{M}) + o_p(1)$ (New Result!) $\tilde{\mathbf{B}}_n$ = -2 log p (\mathcal{D}_n | M]
 $\tilde{\mathbf{B}}_n$ + TRACE $((\tilde{\mathbf{A}}_n)^{-1})$ 1 $g\left(\frac{n}{2\pi}\right) + \log \det \left(\frac{n}{2\pi}\right) - \log \det \left(\frac{n}{2\pi}\right) - \log \det \left(\frac{n}{2\pi}\right)$ $\tilde{l}_n(\hat{\theta}_n; \mathcal{M}) - 2\log p_\theta(\hat{\theta}_n | \mathcal{M}) + q \log \left(\frac{n}{2\pi}\right) -$
 $n\tilde{l}_n(\hat{\theta}_n; \mathcal{M}) - 2\log p_\theta(\hat{\theta}_n | \mathcal{M}) + q \log \left(\frac{n}{2\pi}\right)$
 $2\log p(\mathcal{D}_n | \mathcal{M}) + o_p(1)$ (Lv and Liu, 2014 =2 $n\tilde{l}_n(\hat{\theta}_n; \mathcal{M})$ - 2log $p_\theta(\hat{\theta}_n | \mathcal{M})$ + $q \log\left(\frac{n}{2\pi}\right)$ - log det

= -2log $p(\mathcal{D}_n | \mathcal{M})$ + $o_p(1)$ (Lv and Liu, 2014)

=2 $n\tilde{l}_n(\hat{\theta}_n; \mathcal{M})$ - 2log $p_\theta(\hat{\theta}_n | \mathcal{M})$ + $q \log\left(\frac{n}{2\pi}\right)$ + log det 2 $-2\log p(\mathcal{D}_n \mid \mathcal{M}) + o_p(1)$ (Lv and Liu,
 $2n\tilde{l}_n(\hat{\theta}_n; \mathcal{M}) - 2\log p_\theta(\hat{\theta}_n \mid \mathcal{M}) + q\log\left(\frac{2\log p(\mathcal{D}_n \mid \mathcal{M})}{\epsilon}\right)$ $\left(\mathbf{B}_n\right) = -2\log p\left(\mathbf{\mathcal{D}}_n \mid \mathbf{\mathcal{M}}\right) +$
 $\left(\mathbf{B}_n\right)^{-1} \tilde{\mathbf{B}}_n + TRACE \left(\left(\tilde{\mathbf{A}}_n\right)^{-1} \tilde{\mathbf{B}}_n\right)$ \mathbf{u}) - 2 log $p_{\theta}(\hat{\theta}_n | \mathbf{M}) + q \log \left(\frac{n}{2n} \right)$
 $p_n | \mathbf{M}) + o_p(1)$ (Lv and Liu, 20 = -2 log $p(\mathcal{D}_n | \mathcal{M}) + o_p(1)$ (Lv and Liu, 2014)
 $X = 2n\tilde{l}_n(\hat{\theta}_n; \mathcal{M}) - 2\log p_\theta(\hat{\theta}_n | \mathcal{M}) + q\log\left(\frac{n}{2\pi}\right) + \log \det \tilde{A}_n + TRACE\left((\tilde{A}_n)^{-1} \tilde{B}_n\right)$ $\binom{n}{n} - 2\log p$
*i*_n | *M*) + o_p *TRACE n* $n\tilde{l}_n(\hat{\theta}_n; \mathcal{M}) - 2\log p_\theta(\hat{\theta}_n | \mathcal{M}) + q\log\left(\frac{n}{2\pi}\right) - \log \det\left(\left(\tilde{\mathbf{A}}_n\right)^{-1} \tilde{\mathbf{B}}_n\right) +$
 $2\log p(\mathcal{D}_n | \mathcal{M}) + o_p(1)$ (Lv and Liu, 2014)
 $n\tilde{l}_n(\hat{\theta}_n; \mathcal{M}) - 2\log p_\theta(\hat{\theta}_n | \mathcal{M}) + q\log\left(\frac{n}{2\pi}\right) + \log \det \tilde{\mathbf{A}}$ $p(\mathcal{D}_n | \mathcal{M}) + c$
 $\hat{\theta}_n$; \mathcal{M}) – 2 log
 $p(\ddot{\mathcal{D}}_n | \mathcal{M}) + o$ θ π π $\tilde{\mathbf{B}}_n$ = -2 log p ($\mathcal{D}_n | \mathcal{M}$) + o_p (1) (- $\left(\frac{n}{2\pi}\right)$ + log det $\tilde{A}_n = -2\log p(\mathcal{D}_n \mid \mathcal{M}) + o_p(1)$ (
 $\left(\frac{n}{2\pi}\right)$ - log det $\left(\left(\tilde{A}_n\right)^{-1} \tilde{B}_n\right)$ = -2log $p(\mathcal{D}_n \mid \mathcal{M})$ +
 $\left(\frac{n}{2\pi}\right)$ - log det $\left(\left(\tilde{A}_n\right)^{-1} \tilde{B}_n\right)$ + TRACE $\left(\left(\tilde{A}_n\$ =2n $\tilde{l}_n(\hat{\theta}_n; \mathcal{M})$ - 2log $p_\theta(\hat{\theta}_n | \mathcal{M})$ + $q \log \left(\frac{1}{2} \sum_{n=1}^{\infty} \tilde{l}_n(\hat{\theta}_n; \mathcal{M}) - 2 \log p_\theta(\hat{\theta}_n | \mathcal{M}) + q \log \left(\frac{1}{2} \sum_{n=1}^{\infty} \tilde{l}_n(\hat{\theta}_n | \mathcal{M}) + o_p(\hat{1}) \right)$ (Ly and Liu $\left(\frac{n}{2\pi}\right) - \log \det \left(\left(\tilde{\mathbf{A}}_n\right)^{-1} \tilde{\mathbf{B}}_n\right)$
2014)
 $\left(\frac{n}{2\pi}\right) + \log \det \tilde{\mathbf{A}}_n + TRA$ $-2\log p_{\theta}\left(\hat{\theta}_{n} | \mathbf{M}\right) + q \log \left(\frac{n}{2\pi}\right) - \log \det \left(\left(\tilde{\mathbf{A}}_{n}\right)^{-1} \tilde{\mathbf{B}}_{n}\right) + TRACE\left(\left(\mathbf{M}\right) + o_{p}\left(1\right) \right)$ (Lv and Liu, 2014)
 $-2\log p_{\theta}\left(\hat{\theta}_{n} | \mathbf{M}\right) + q \log \left(\frac{n}{2\pi}\right) + \log \det \tilde{\mathbf{A}}_{n} + TRACE\left(\left(\tilde{\mathbf{A}}_{n}\right)^{-1} \tilde$ = $-2\log p(\mathcal{D}_n | \mathcal{M}) + o_p(1)$ (Lv and Lit

x = $2n\tilde{l}_n(\hat{\theta}_n; \mathcal{M}) - 2\log p_\theta(\hat{\theta}_n | \mathcal{M}) + q\log$

= $-2\log p(\mathcal{D}_n | \mathcal{M}) + o_p(1)$ (New Result $\begin{aligned} &\mathbf{A}_n = -2\log p\left(\mathbf{\mathcal{D}}_n \mid \mathbf{\mathcal{M}}\right) + o_p\left(1\right) \ &\mathbf{A}_n\big)^{-1} \mathbf{\tilde{B}}_n\bigg) = -2\log p\left(\mathbf{\mathcal{D}}_n \mid \mathbf{\mathcal{M}}\right) + \ &\mathbf{\tilde{A}}_n\bigg)^{-1} \mathbf{\tilde{B}}_n\bigg) + TRACE\bigg(\bigg(\mathbf{\tilde{A}}_n\bigg)^{-1} \mathbf{\tilde{B}}_n\bigg) \end{aligned}$ **GBIC**_P = $2n\tilde{l}_n(\hat{\theta}_n; \mathcal{M}) - 2\log p_\theta(\hat{\theta}_n | \mathcal{M}) + q\log\left(\frac{n}{2\pi}\right) - \log \det((\tilde{A}_n)^{-1}\tilde{B}_n) + TRACE((\tilde{A}_n)^{-1}\tilde{B}_n)$

= $-2\log p(\mathcal{D}_n | \mathcal{M}) + o_p(1)$ (Lv and Liu, 2014)
 GBIC_X = $2n\tilde{l}_n(\hat{\theta}_n; \mathcal{M}) - 2\log p_\theta(\hat{\theta}_n | \math$ **a** $(-2)^n$
D $(-2)^n$
D $(-2)^n$ M) - 2 $\log p_{\theta}(\hat{\theta}_n \mid M)$
 $\mathcal{D}_n \mid M$) + $o_p(1)$ (Lv
 M) - 2 $\log p_{\theta}(\hat{\theta}_n \mid M)$ $\left(\begin{array}{c|c} \mathbf{\mathcal{D}}_{\!n}\mid\mathbf{\mathcal{M}}\end{array}\right)$

ABSTRACT

Bayesian model selection criteria (BMSC) require the evaluation of a computationally intractable multidimensional integral. Although computationally expensive Monte Carlo simulation methods may be used for such evaluations, Laplace approximation methods provide a computationally inexpensive alternative approach. In this paper, a computationally intractable multidimensional BMSC integral is approximated using a Laplace approximation to obtain a new BMSC called GBIC_X . With respect to seven real world data sets, GBIC_x exhibited performance which was superior to BIC-family model selection criteria for AIC-biased simulation studies and showed performance which was superior to AIC-family model selection criteria for BICbiased simulation studies. These findings suggest that $GBIC_x$ may be especially useful in situations where a more robust BMSC approximation is desirable. **Theory Simulations**

$$
\mathcal{M}\big)
$$

$$
\log p\left(\mathbf{x}_i | \hat{\boldsymbol{\theta}}_n, \mathcal{M}\right)
$$
\n
$$
il\left(\boldsymbol{\theta}; \mathcal{M}\right)
$$
\n
$$
p_{\theta}(\boldsymbol{\theta} | \mathcal{M})d\boldsymbol{\theta} \text{ with}
$$

$$
P_{\theta}(\mathbf{v}|\mathbf{v}\mathbf{v})
$$
 with prior $P_{\theta}(\mathbf{v}|\mathbf{v}\mathbf{v})$

(1) only if
$$
p_o \in \mathcal{M}
$$
 (Akaike, 1974)

$$
\mathbf{g} \left[\mathbf{p} \left(\mathbf{\tilde{\mathbf{D}}}_{n} \mid \mathbf{\theta}^{*}, \mathbf{\mathcal{M}} \right) \right] + o_{p} \left(1 \right) \ \left(\text{Takeuchi, 1976} \right)
$$

Theorem 2.1 (GBIC Cross-Entropy Approximation). Assume Assumptions $A1 - A6$ hold. Let the model prior probability density $p_{\theta}: \Theta \to [0,\infty)$ be a continous function on Θ such that for all $\theta \in \Theta$: $p_{\theta}(\theta) > 0$. Let $p(\mathcal{D}_n|\mathcal{M}) \equiv exp(-n\ell(\theta))$ where $\ell(\theta) \equiv -\int p_o(\mathbf{x}) \log p(\mathbf{x}|\theta) d\nu(\mathbf{x}) < \infty$. Assume there exists a number n_0 such that for all $n \geq n_0$: $p(\mathcal{D}_n|\mathcal{M}) < \infty$. Then as $n \to \infty$,

 $-(1/n)\log p(\ddot{\mathcal{D}}_n|\mathcal{M})=E\{\tilde{\ell}_n(\hat{\boldsymbol{\theta}}_n)\}+(1/(2n))$

 $\frac{\log p_\theta(\hat{\boldsymbol\theta}_n|\mathcal{M})}{n} + \frac{q}{2n}\log\left(\frac{n}{2\pi}\right) -$

Proof. First, use the Multidimensional Laplace Approximation Theorem ([2], pp. 86-88) leaving $\ell(\theta^*)$, \mathbf{A}^* , \mathbf{B}^* , and $p_{\theta}(\theta^*)$ to be estimated. The estimators $\mathbf{\hat{A}}_n = \mathbf{A}^* + o_p(1)$, $\mathbf{\hat{B}}_n = \mathbf{B}^* + o_p(1)$, and $p_{\theta}(\hat{\theta}_n) = p_{\theta}(\theta^*) + o_p(1)$ can be substituted to estimate \mathbf{A}^* , \mathbf{B}^* , and $p_{\theta}(\theta^*)$ respectively because in conjunction with the existing assumptions the resulting approximation error associated with these substitutions in (4) is $o_p(1/n)$. Second, Proposition P2 of Linhart and Volkers (1984)(see [9]) shows that

 $\ell(\boldsymbol{\theta}^*) = E\{\tilde{\ell}_n(\hat{\boldsymbol{\theta}}_n)\} + (1/(2n))$

ensure the approximation error in (4) is $o_p(1/n)$.

Veural Information Processing Systems

$$
))TRACE\left[({\bf \hat A}_n)^{-1}{\bf \hat B}_n\right]-
$$

$$
\frac{n}{2\pi}\Big) + \frac{\log(\det(\hat{\mathbf{A}}_n))}{2n} + o_p\left(\frac{1}{n}\right). \tag{4}
$$

$$
a))TRACE\left[(\hat{\mathbf{A}}_n)^{-1}\hat{\mathbf{B}}_n\right] + o_p(1/n). \tag{5}
$$

Thus, Equation (5) must be used rather than $\ell(\theta^*) = E\{\tilde{\ell}_n(\hat{\theta}_n)\} + O_p(1/n)$ to estimate $\ell(\theta^*)$ to