# Bayesian Model Updating and Uncertainty Quantification: Theory, Computational Tools, and Applications

Babak Moaveni, Tufts University

Costas Papadimitriou, University of Thessaly

# Background

- Thomas Bayes (1701 – 1761) was an English statistician, philosopher and minister.
- Bayes most famous paper including what is known as Bayes Theorem was published in 1763, two years after his death by Richard Price.
- Bayes Theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

# Background

**Bayesian vs. Frequentist Interpretation of Probability**

- Bayesian or epistemological: probability measures a degree of belief. Bayes theorem then links the degree of belief in a proposition before and after accounting for evidence.
- Frequentist: probability measures a *proportion of outcomes*.

# Bayesian Inference

- Assume model $\mathcal{M}$ is characterized by parameters $\boldsymbol{\theta}$
- Bayes Theorem: $$p(\boldsymbol{\theta}\,|\,\mathbf{d}) = \frac{p(\mathbf{d}\,|\,\boldsymbol{\theta})\,p(\boldsymbol{\theta})}{p(\mathbf{d})} = c\;p(\mathbf{d}\,|\,\boldsymbol{\theta})\,p(\boldsymbol{\theta})$$

$\mathbf{d}$: Measured data

$\boldsymbol{\theta}$: Model parameter

$p(\boldsymbol{\theta}|\mathbf{d})$: Posterior conditional probability of $\boldsymbol{\theta}$ given $\mathbf{d}$

$p(\mathbf{d}|\boldsymbol{\theta})$: Likelihood function

$p(\boldsymbol{\theta})$: Prior probability distribution of $\boldsymbol{\theta}$

$p(\mathbf{d})$: evidence

$c$: Normalization constant

# Bayesian Inference

***<u>Example 1 (Yuen 2010):</u>*** Estimate the mean and variance of a Gaussian random variable X from N measurements.

$\mathbf{d} = \{x_1, x_2, \ldots, x_N\}$; independent measurements

$\boldsymbol{\theta} = \{\mu, \sigma^2\}$

# Bayesian Inference

Assuming an uninformative prior →

$$p(\boldsymbol{\theta} \mid \mathbf{d}) = \frac{c}{\sigma^N} \exp\left[ -\frac{1}{2\sigma^2} \sum_{n=1}^{N} (x_n - \mu)^2 \right]$$

Negative log-likelihood

$$J(\boldsymbol{\theta} \mid \mathbf{d}) = N \ln \sigma + \frac{1}{2\sigma^2} \sum_{n=1}^{N} (x_n - \mu)^2 + constant$$

The most probable estimate for μ:

$$\partial J / \partial \mu = 0 \Rightarrow \mu^* = \frac{1}{N} \sum_{n=1}^{N} x_n$$

Conditional posterior PDF of $p(\mu \mid \sigma^2, \mathbf{d})$ is a Gaussian distribution with mean $\mu^*$ and variance $\sigma^2/N$.

# Bayesian Inference

Note that variance of estimated mean decreases as more data points are collected.

Similarly, the most probable estimate for $\sigma^2$:

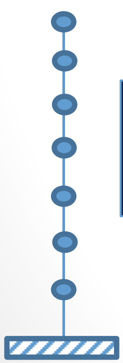
$$\partial J \big/ \partial \sigma^2 = 0 \Rightarrow \sigma^{2*} = \frac{1}{N} \sum_{n=1}^{N} (x_n - \mu^*)^2$$

Conditional posterior PDF of $p\left(\sigma^2 \mid \mu, \mathbf{d}\right)$ is a Gamma distribution.

# Model Updating

Measured data features

Optimal model par θ to minimize difference between model and data



**Dynamic Model $M(\theta)$**

Model-predicted features

# Model Updating

- Model updating can be deterministic or probabilistic
- In deterministic approach, model parameters $\boldsymbol{\theta}$ are updated by minimizing an objective function defined as

$$J(\boldsymbol{\theta}) = \mathbf{e}(\boldsymbol{\theta})^T \mathbf{W} \mathbf{e}(\boldsymbol{\theta}) = \sum_{j=1}^{N_r} w_j e_j^{\,2}(\boldsymbol{\theta})$$

- Residual $e_j$ correspond to the difference between model-predicted and measured data features
- Most common data features: input-output vibration data, modal parameters
- Optimal parameters are found using local optimization methods (e.g., Gauss-Newton) or global methods (e.g., genetic algorithm, simulated annealing).

# Model Updating

**Challenges**

- ill-conditioning of the inverse problem due to:
  - Insufficient information from the measured data
  - "Unidentifiability" when having large number of updating parameters
- Effects of modeling errors on prediction accuracy of updated models
- High computational cost for global optimization
- Sensitivity of results to assigned weights and residuals
- Effects of changing ambient conditions on modeling assumptions

# Bayesian Model Updating

- Bayes theorem: $p(\boldsymbol{\theta}|\mathbf{d}) \propto p(\mathbf{d}|\boldsymbol{\theta})\,p(\boldsymbol{\theta})$
- For data feature $\mathbf{d}$, the error between measured and model predicted data is $\mathbf{e_d} = \tilde{\mathbf{d}} - \mathbf{d}(\boldsymbol{\theta})$
- The probability distribution of $\mathbf{e}$ is assumed based on Maximum Information Entropy which results in a (zero mean) Gaussian distribution*.

$$p\left(\tilde{d}_i \,|\, \boldsymbol{\theta}\right) = p\left(e_{d_i} \,|\, \boldsymbol{\theta}\right) = \frac{1}{\sqrt{2\pi}\sigma_{d_i}} \exp\left(-\frac{1}{2}\frac{e_{d_i}^2}{\sigma_{d_i}^2}\right)$$

*Any bias can be added to* $\mathbf{d}(\boldsymbol{\theta})$ → *zero mean*

# Bayesian Model Updating

- Assuming independent measurements:

$$p(\mathbf{d}\,|\,\boldsymbol{\theta}) = \prod_{i=1}^{N_i} p(\tilde{d}_i\,|\,\boldsymbol{\theta})$$

- Common data features are response time histories (acceleration) and modal parameters (natural freq., mode shapes)
- For the case of $d_i$ being vector of response time history with $N_0$ points at sensor $i$ ($\mathbf{y}_i$) and covariance matrix $\boldsymbol{\Sigma}$: *

$$p(\mathbf{d}\,|\,\boldsymbol{\theta}) = \frac{1}{(2\pi)^{N_0 N_i/2}\left\|\boldsymbol{\Sigma}\right\|^{1/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{N_i}\left(\tilde{\mathbf{y}}_i - \mathbf{y}_i(\boldsymbol{\theta})\right)^T \boldsymbol{\Sigma}^{-1}\left(\tilde{\mathbf{y}}_i - \mathbf{y}_i(\boldsymbol{\theta})\right)\right)$$

** Assuming similar covariance matrix for different sensors*

# Bayesian Model Updating

- In case of having a non-informative prior (e.g., uniform),

$$p(\boldsymbol{\theta} \mid \mathbf{d}) = p(\mathbf{d} \mid \boldsymbol{\theta})$$

- Thus, the maximum a-postriori (MAP) are the same as maximum likelihood (ML) estimates which can be the same as deterministic optimal parameters

$$\boldsymbol{\theta}^{MAP} = \boldsymbol{\theta}^{ML} = \arg\max\{p(\mathbf{d} \mid \boldsymbol{\theta})\} = \arg\min\{-\log(p(\mathbf{d} \mid \boldsymbol{\theta}))\}$$

$$J(\boldsymbol{\theta}) = -\log(p(\mathbf{d} \mid \boldsymbol{\theta})) = \frac{1}{2}\sum_{i=1}^{N_i}\left(\tilde{\mathbf{y}}_i - \mathbf{y}_i(\boldsymbol{\theta})\right)^T \boldsymbol{\Sigma}^{-1}\left(\tilde{\mathbf{y}}_i - \mathbf{y}_i(\boldsymbol{\theta})\right) + \frac{1}{2}\log((2\pi)^{N_i N_0}\|\boldsymbol{\Sigma}\|)$$

$$= \frac{1}{2}\sum_{i=1}^{N_i}\mathbf{e}(\boldsymbol{\theta})_i^{\,T}\,\mathbf{W}\mathbf{e}(\boldsymbol{\theta})_i + constant$$

# Bayesian Model Updating

- For the case of $d_i$ corresponding to the $i^{\text{th}}$ mode eigenfrequency (square natural freq.) and mode shape from $N_i$ sensors

$$\mathbf{d}_m = \{\tilde{\lambda}_m, \tilde{\mathbf{\Phi}}_m\}$$

- The error function can be written as:

$$\tilde{\lambda}_m - \lambda_m(\boldsymbol{\theta}) = e_{\lambda_m} \sim N\left(0, \sigma^2_{\lambda_m}\right)$$

$$\tilde{\mathbf{\Phi}}_m - a_m \mathbf{\Phi}_m(\boldsymbol{\theta}) = \mathbf{e}_{\mathbf{\Phi}_m} \sim N\left(\mathbf{0}, \mathbf{\Sigma}_{\mathbf{\Phi}_m}\right); \qquad a_m = \frac{\tilde{\mathbf{\Phi}}_m^T . \mathbf{\Gamma}\mathbf{\Phi}_m(\boldsymbol{\theta})}{\left\|\mathbf{\Gamma}\mathbf{\Phi}_m(\boldsymbol{\theta})\right\|^2}$$

# Bayesian Model Updating

- Assuming independent measurements:

$$p(\mathbf{d}\,|\,\boldsymbol{\theta}) = p\left(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\Phi}}\,|\,\boldsymbol{\theta}\right) = \prod_{m=1}^{N_m} p\left(\tilde{\lambda}_m\,|\,\boldsymbol{\theta}, \sigma_{\lambda_m}^2\right) p\left(\tilde{\boldsymbol{\Phi}}_m\,|\,\boldsymbol{\theta}, \boldsymbol{\Sigma}_{\boldsymbol{\Phi}_m}\right)$$

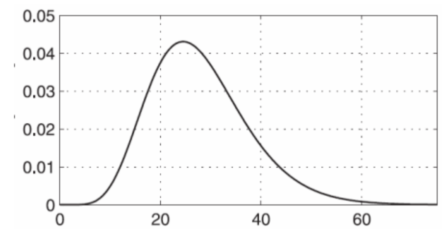
$$p(\mathbf{d}\,|\,\boldsymbol{\theta}) \propto \exp\left(-J(\boldsymbol{\theta})\right)$$

$$J(\boldsymbol{\theta}) = \sum_{m=1}^{N_m} \frac{1}{2\sigma_{\lambda_m}^2}\left(\tilde{\lambda}_m - \lambda_m(\boldsymbol{\theta})\right)^2 + \sum_{m=1}^{N_m} \frac{1}{2\sigma_{\boldsymbol{\Phi}_m}^2}\left(\tilde{\boldsymbol{\Phi}}_m - a_m \boldsymbol{\Phi}_m(\boldsymbol{\theta})\right)^T \left(\tilde{\boldsymbol{\Phi}}_m - a_m \boldsymbol{\Phi}_m(\boldsymbol{\theta})\right)$$

- The maximum likelihood (ML) estimate can be reached by minimizing $-\log\left(p\left(\mathbf{d}\,|\,\boldsymbol{\theta}\right)\right)$ or $J(\boldsymbol{\theta})$ which is similar to the deterministic optimization approach.

# Bayesian Model Updating

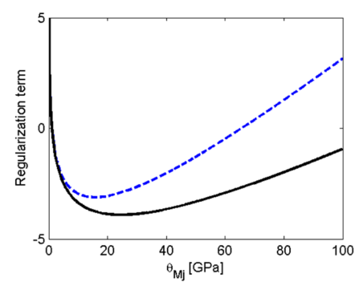
- In the case of having an informative (non-uniform) prior:
  maximum a-posteriori (MAP) estimate ≠ ML estimate
- Effects of prior distribution in Bayesian model updating is similar to effects of regularization term in deterministic model updating
- Assume prior = joint Gamma distribution




$$p(\boldsymbol{\theta}) = \prod_{i=1}^{N_\theta} \frac{\theta_i^{\alpha_i - 1}}{\beta_i^{\alpha_i}\Gamma(\alpha_i)} \exp\left(-\frac{\theta_i}{\beta_i}\right)$$

$$p(\boldsymbol{\theta} \mid \mathbf{d}) \propto p(\mathbf{d} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta})$$

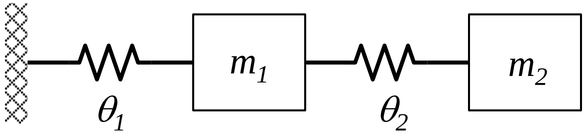
$$p(\boldsymbol{\theta} \mid \mathbf{d}) \propto \exp\left(-J_{ML}(\boldsymbol{\theta}) - \sum_i^{N_\theta}\left(\frac{\theta_i}{\beta_i} + (1-\alpha_i)\log\theta_i\right)\right)$$




$$J_{MAP} = J_{ML} + \sum_i^{N_\theta}\left(\frac{\theta_i}{\beta_i} + (1-\alpha_i)\log\theta_i\right)$$

# Example

- ***<u>Example 2 (Yuen 2010):</u>*** Consider in the two-DOF system:




masses are known as $m_1 = m_2 = 1$ and $\theta_1$, $\theta_2$ are to be estimated

Classify the identifiability of this system:

- CASE 1- Based on one measured eigenvalue
- CASE 2- Based on two measured eigenvalues
- CASE 3- Based on one set of eigenvalue and mode shape (eigenvector)

# Example

- Assume the true values of stiffness values are $\theta_1 = \theta_2 = 1$

$$\mathbf{K} = \begin{bmatrix} \theta_1 + \theta_2 & -\theta_2 \\ -\theta_2 & \theta_2 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}, \mathbf{M} = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{K}\phi_i = \lambda_i \mathbf{M}\phi_i \qquad i = 1, 2$$

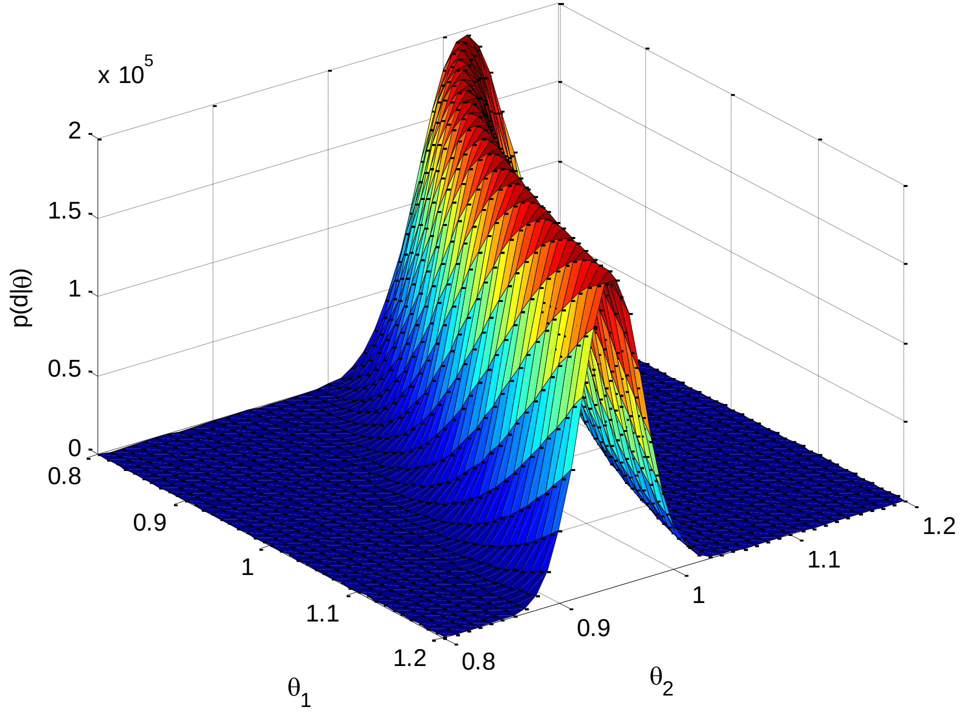
$$\Rightarrow \lambda_1 = 0.382,\ \lambda_2 = 2.618,\ \phi_1 = \begin{bmatrix} 1 \\ 1.618 \end{bmatrix},\ \phi_2 = \begin{bmatrix} 1 \\ -0.618 \end{bmatrix}$$

- CASE 1: Assume $\lambda_1$ is measured with 5% Gaussian noise and five measurements are available: $\mathbf{d}_1 = \{0.3860, 0.3922, 0.4157, 0.3592, 0.3615\}$

- Assume a known $\sigma_{\lambda_1} = 0.05 \times 0.382 = 0.0191$

# Example

- To calculate $\lambda_1(\theta) \rightarrow \left|\mathbf{K} - \lambda_1 \mathbf{M}\right| = 0 \;\Rightarrow\; \lambda_1{}^2 - (\theta_1 + 2\theta_2)\lambda_1 + \theta_1\theta_2 = 0$
- No unique optimal parameters → unidentifiable

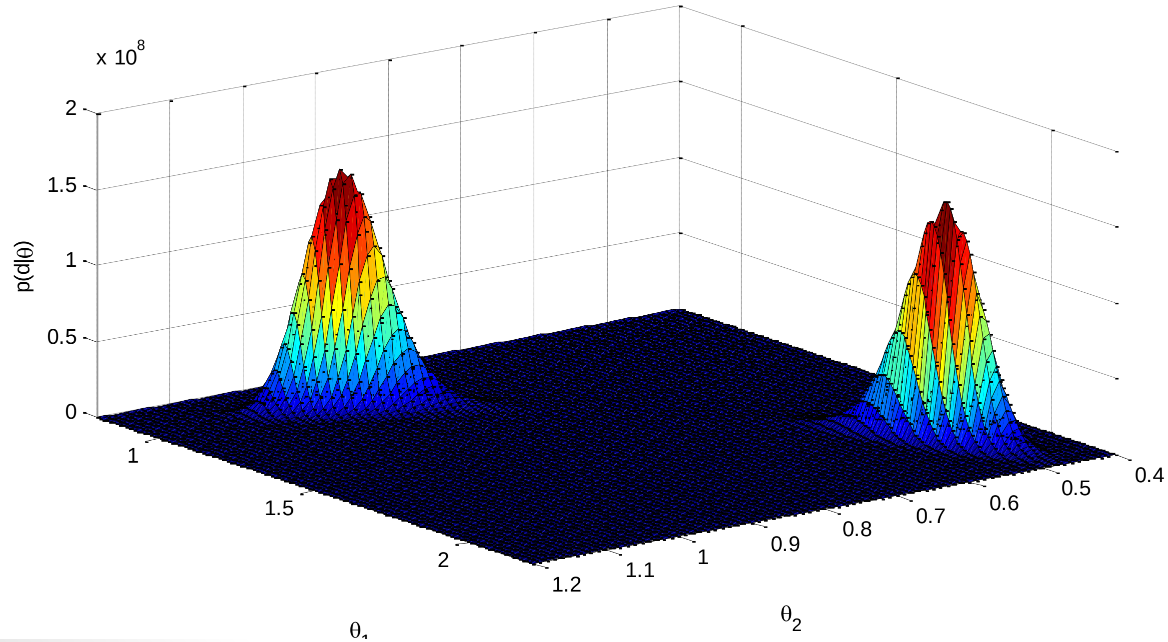


# Example

- CASE 2: Assume both eigenvalues are measured with 5% Gaussian noise

$$\boldsymbol{\lambda}_2 = \{2.3614, 2.5877, 2.7070, 2.3875, 2.7272\}$$

$$p(\mathbf{d}\,|\,\boldsymbol{\theta}) = \frac{1}{(\sqrt{2\pi})^{10}\sigma_{\lambda_1}^5\sigma_{\lambda_2}^5}\exp\left[-\frac{1}{2\sigma_{\lambda_1}^2}\sum_{m=1}^{5}\left(\tilde{\lambda}_{1m}-\lambda_1(\boldsymbol{\theta})\right)^2-\frac{1}{2\sigma_{\lambda_2}^2}\sum_{m=1}^{5}\left(\tilde{\lambda}_{2m}-\lambda_2(\boldsymbol{\theta})\right)^2\right]$$

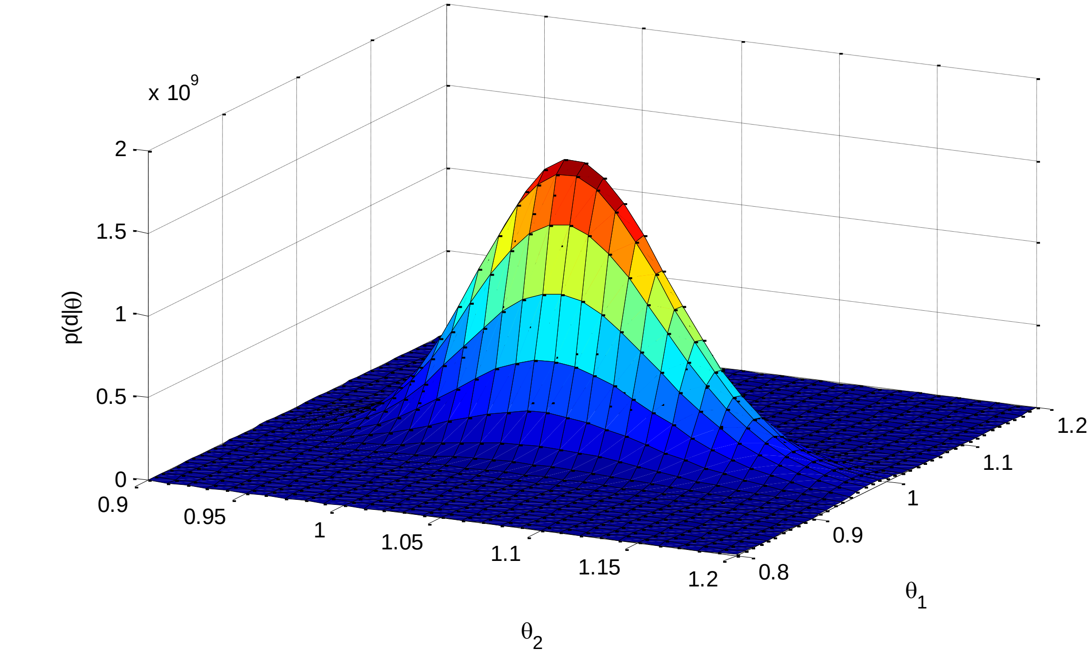



Locally Identifiable

# Example

- CASE 3: Assume first eigenvalue and mode shape are measured with 5% Gaussian noise

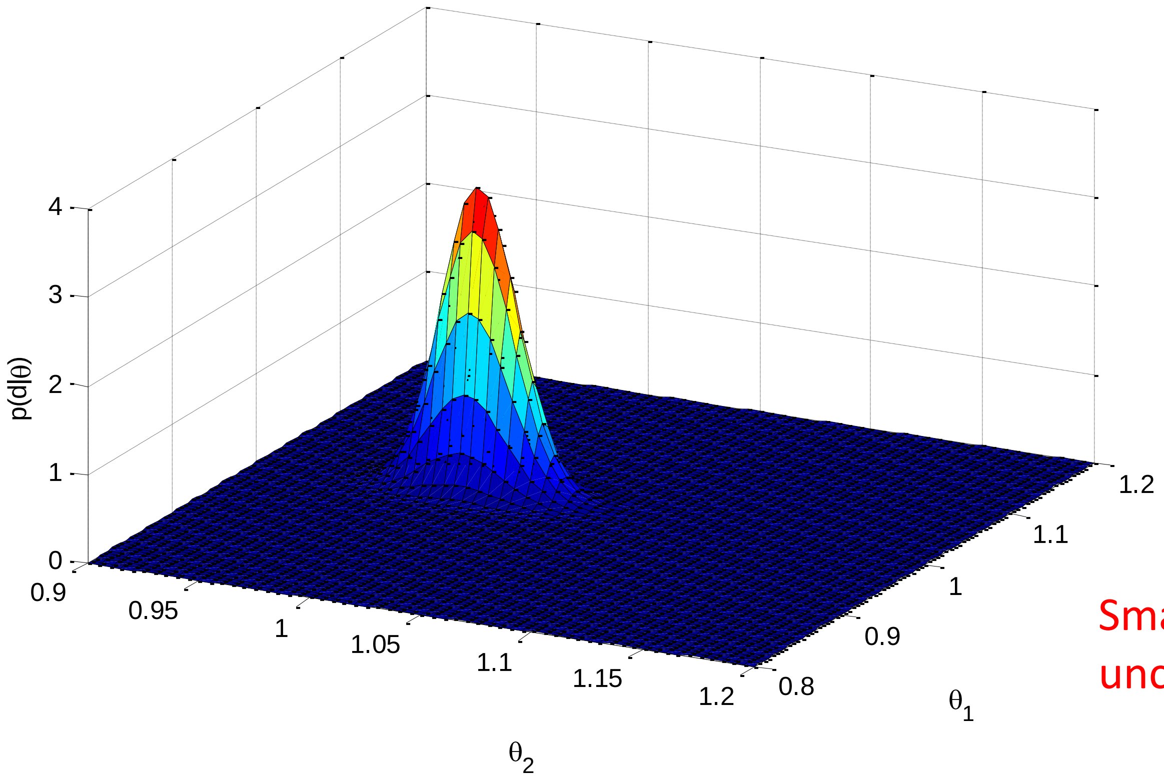
$$p(\mathbf{d}\,|\,\boldsymbol{\theta}) = \frac{1}{(\sqrt{2\pi})^{10}\sigma_{\lambda_1}^5\sigma_{\phi}^5}\exp\left[-\frac{1}{2\sigma_{\lambda_1}^2}\sum_{m=1}^{5}\left(\tilde{\lambda}_{1m}-\lambda_1(\boldsymbol{\theta})\right)^2-\frac{1}{2\sigma_{\phi}^2}\sum_{m=1}^{5}\left(\frac{\tilde{\phi}_{12}}{\tilde{\phi}_{11}}-\frac{\phi_{12}(\boldsymbol{\theta})}{\phi_{11}(\boldsymbol{\theta})}\right)^2\right]$$




Globally Identifiable

# Example

- Now if instead of 5 data points, we have 30 measured data points in CASE 3:




Smaller estimation uncertainty for $\theta_1$ and $\theta_2$

# References


- K.V. Yuen, Bayesian methods for structural dynamics and civil engineering, Wiley Singapore, 2010.
- J.L. Beck, Bayesian system identification based on probability logic, Structural Control and Health Monitoring, 17 (2010) 825-847.
- J.L. Beck, L.S. Katafygiotis, Updating models and their uncertainties. I: Bayesian statistical framework, Journal of Engineering Mechanics, ASCE, 124 (1998) 455-461.