Bayesian Defect Signal Analysis for Nondestructive Evaluation of Materials†

Aleksandar Dogandžić
Electrical and Computer Engineering
Iowa State University

†Joint work with Benhong Zhang (Ph.D. student), supported by the NSF I-U Cooperative Research Program, CNDE, Iowa State University.

My Research Topics

General Area: Statistical Signal Processing.

Applications include:

- Sensor networks,
- Nondestructive evaluation and testing (NDE/NDT) of materials,
- Wireless communications,
- Radar.

Background: What is NDE?

In nondestructive evaluation (NDE) of materials, *noninvasive measurement techniques* are used to determine if various properties of materials, components, or structures have defects that might lessen their ability to perform their intended function.

Background: What is NDE?

In nondestructive evaluation (NDE) of materials, *noninvasive measurement techniques* are used to determine if various properties of materials, components, or structures have defects that might lessen their ability to perform their intended function.

Typical goals of NDE inspections:

- determining if the wing of an airliner has corrosion around the rivets or
- testing materials for flaws that might cause failure (before the material is used to make parts or components).

Our Goal in This Talk

- **Goal:** Find elliptically-shaped regions with elevated signal levels in noisy images.

- **Application:** Automatic identification of defects
  (e.g. cracks, corrosion, porosity, inclusions etc).

Accounting for both the defect-signal amplitude (signal level) and area greatly improves the detectability performance (compared with the traditional NDE methods which ignore the spatial extent of the defect signals).

Approach

Statistical Modeling Approach: Hierarchical Bayesian.

Computational Approach: Markov chain Monte Carlo, for simulating from the posterior distributions of the desired parameters $\Longrightarrow$ Intuitively, generate a bunch of "proposals" for the defect regions and signals; if a proposal matches the data well, generate more "similar" proposals.

Outline

- Measurement Model
  - Parametric model for defect location and shape,
  - Measurement-error (noise) model,
  - Defect-signal (reflectivity) model,
  - Prior specifications for the location, shape, and defect-signal distribution parameters (model parameters).
- Bayesian Analysis
  - Simulating the model parameters $\boldsymbol{\phi}$,
  - Simulating the signals θ_i.
- Numerical Examples.

Terminology

Abbreviations:

- i.i.d. $\equiv$ independent, identically distributed,
- pdf $\equiv$ probability density function,
- cdf $\equiv$ cumulative distribution function,
- MMSE $\equiv$ minimum mean-square error,
- MCMC $\equiv$ Markov Chain Monte Carlo.

Notation

- "T" $\equiv$ a transpose,
- $\pi_{\boldsymbol{\phi}}(\boldsymbol{\phi}) \equiv$ prior pdf of $\boldsymbol{\phi}$,
- $p(\boldsymbol{\phi} \,|\, \boldsymbol{y}) \equiv$ conditional pdf of $\boldsymbol{\phi}$ given $\boldsymbol{y}$,
- $\mathcal{R}(\boldsymbol{z}) \equiv$ defect region,
- $\mathcal{R}^{\mathrm{c}}(\boldsymbol{z}) \equiv$ noise-only region, i.e. region outside $\mathcal{R}(\boldsymbol{z})$.
- $$i_A(x) = \begin{cases} 1, & x \in A, \\ 0, & \text{otherwise} \end{cases} \quad \equiv \quad \text{indicator function.}$$

Notation (cont.)

- $\mathcal{N}(y;\mu,\sigma) \equiv$ Gaussian pdf of a random variable y with mean μ and standard deviation σ:

$$\mathcal{N}(y;\,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right],$$

- $\mathcal{N}_{\mathrm{t}}(y\,;\,\mu,\sigma) \equiv$ truncated-Gaussian pdf (to non-negative values) with parameters μ and σ:

$$\mathcal{N}_{\mathrm{t}}(y\,;\,\mu,\sigma) = \frac{\mathcal{N}(y\,;\,\mu,\sigma)}{\Phi(\mu/\sigma)} \cdot i_{[0,\infty)}(y)$$

where $\Phi(\cdot)$ denotes the cdf of the standard normal random variable.

Outline

- Measurement Model
  - Parametric model for defect location and shape,
  - Measurement-error (noise) model,
  - Defect-signal (reflectivity) model,
  - Prior specifications for the location, shape, and defect-signal distribution parameters (model parameters).

- Bayesian Analysis
  - Simulating the model parameters ϕ,
  - Simulating the signals θ_i.

- Numerical Examples.

Parametric Model for Defect Location and Shape

Potential defect-signal region modeled as an ellipse $\mathcal{R}(\boldsymbol{z})$:

$$\mathcal{R}(\boldsymbol{z}) = \{\boldsymbol{r} : (\boldsymbol{r} - \boldsymbol{r}_0)^T \Sigma_{\mathcal{R}}^{-1} (\boldsymbol{r} - \boldsymbol{r}_0) \le 1\}$$

where

- $\boldsymbol{r} = [x_1, x_2]^T \equiv$ location in Cartesian coordinates
- $\boldsymbol{z} = [\boldsymbol{r}_0^T, d, A, \varphi]^T \equiv$ vector of the (unknown) location and shape parameters, and

$$\Sigma_{\mathcal{R}} = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \cdot \begin{bmatrix} d^2 & 0 \\ 0 & A^2/(d^2\pi^2) \end{bmatrix} \cdot \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}^T.$$

Parametric Model for Defect Location and Shape (cont.)

Ellipse location and shape parameters:

- $\boldsymbol{r}_0 = \begin{bmatrix} x_{0,1} \\ x_{0,2} \end{bmatrix} \equiv$ center of the ellipse in Cartesian coordinates (m);
- $d > 0 \equiv$ axis parameter (m);
- $A \equiv$ area of the ellipse (m^2);
- $\varphi \in [-\pi/4, \pi/4] \equiv$ ellipse orientation parameter (rad).

Parametric Model for Defect Location and Shape (cont.)

Ellipse location and shape parameters:

- $\boldsymbol{r}_0 = \begin{bmatrix} x_{0,1} \\ x_{0,2} \end{bmatrix} \equiv$ center of the ellipse in Cartesian coordinates (m);
- $d > 0 \equiv$ axis parameter (m);
- $A \equiv$ area of the ellipse (m^2);
- $\varphi \in [-\pi/4, \pi/4] \equiv$ ellipse orientation parameter (rad).

Note: d and $A/(d\pi)$ are the axes of this ellipse.

Outline

- Measurement Model
  - Parametric model for defect location and shape,
  - Measurement-error (noise) model,
  - Defect-signal (reflectivity) model,
  - Prior specifications for the location, shape, and defect-signal distribution parameters (model parameters).

- Bayesian Analysis
  - Simulating the model parameters ϕ,
  - Simulating the signals θ_i.

- Numerical Examples.

Measurement-Error (Noise) Model

Assume that we have collected measurements y_i at locations $\boldsymbol{s}_i$, $i = 1, 2, \ldots, \mathbb{N}_{\rm tot}$ within the region of interest, where $\mathbb{N}_{\rm tot}$ denotes the total number of measurements in this region. We adopt the following measurement-error (noise) model:

- If y_i is collected over the defect region [i.e. $\boldsymbol{s}_i \in \mathcal{R}(\boldsymbol{z})$], then

$$y_i = \underbrace{\theta_i}_{\text{signal}} + \underbrace{e_i}_{\text{additive noise}}$$

  where θ_i and e_i denote the defect signal (related to its reflectivity) and noise at location $\boldsymbol{s}_i$, respectively;

Measurement-Error (Noise) Model (cont.)

- If y_i is collected outside the defect region [i.e. $\boldsymbol{s}_i \in \mathcal{R}^{\mathrm{c}}(\boldsymbol{z})$], then

$$y_i = \underbrace{0}_{\text{signal}} + \underbrace{e_i}_{\text{additive noise}}$$

  implying that the signals θ_i are zero in the noise-only region;

Measurement-Error (Noise) Model (cont.)

- If y_i is collected outside the defect region [i.e. $\boldsymbol{s}_i \in \mathcal{R}^{\mathrm{c}}(\boldsymbol{z})$], then

$$y_i = \underbrace{0}_{\text{signal}} + \underbrace{e_i}_{\text{additive noise}}$$

  implying that the signals θ_i are zero in the noise-only region;

- We model the additive noise samples e_i, $i = 1, 2, \ldots, \mathbb{N}_{\text{tot}}$ as zero-mean i.i.d. Gaussian random variables with known variance σ^2 (which can be easily estimated from the noise-only data).

Measurement-Error (Noise) Model (cont.)

Therefore,

$$p(y_i \,|\, \theta_i) = \mathcal{N}(y_i \,;\, \theta_i, \sigma)$$

where $\theta_i = 0$ for $\boldsymbol{s}_i \in \mathcal{R}^{\mathrm{c}}(\boldsymbol{z})$.

Outline

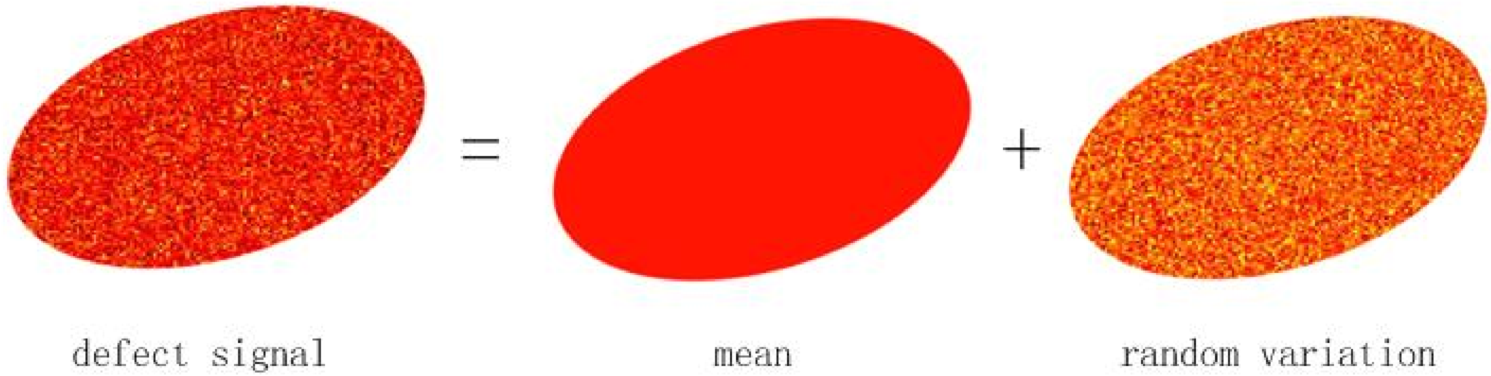
- Measurement Model
  - Parametric model for defect location and shape,
  - Measurement-error (noise) model,
  - Defect-signal (reflectivity) model,
  - Prior specifications for the location, shape, and defect-signal distribution parameters (model parameters).
- Bayesian Analysis
  - Simulating the model parameters $\boldsymbol{\phi}$,
  - Simulating the signals θ_i.
- Numerical Examples.

Defect-Signal (Reflectivity) Model

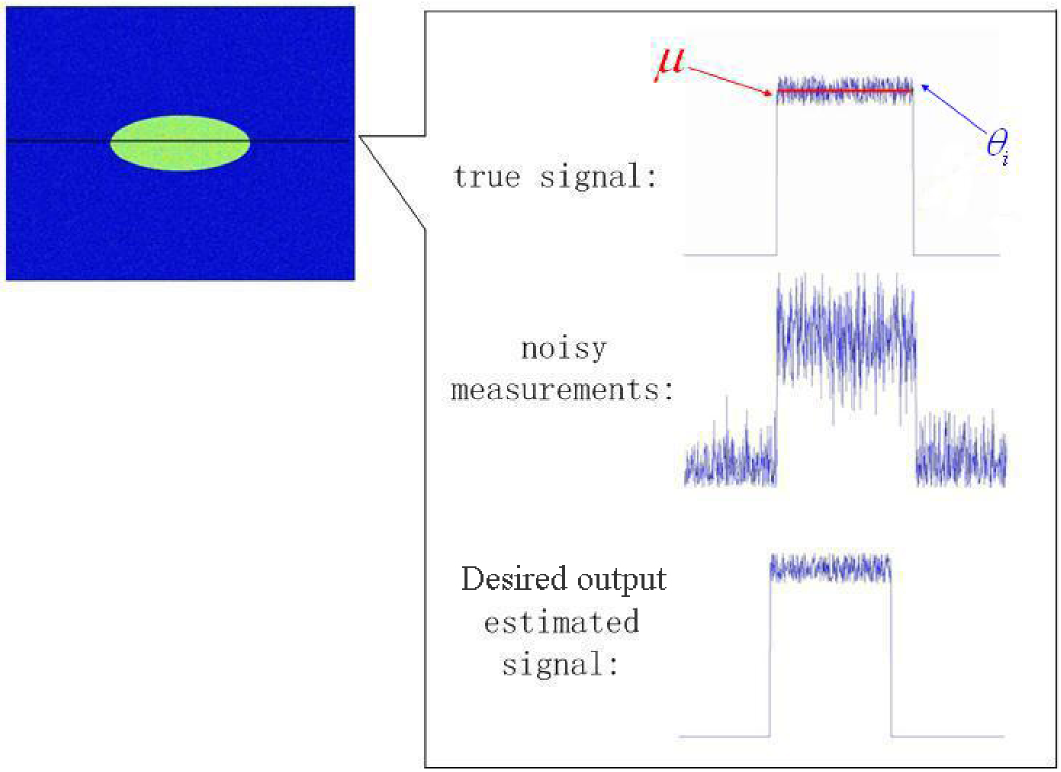
The signals within the defect region [i.e. $\boldsymbol{s}_i \in \mathcal{R}(\boldsymbol{z})$] are i.i.d. truncated Gaussian with unknown *defect-signal distribution parameters* μ *and* τ. Therefore, the joint pdf of the defect signals conditional on $\boldsymbol{z}$ (location and shape), μ, and τ (*defect-signal distribution parameters*) is

$$p(\{\theta_i,\, \boldsymbol{s}_i \in \mathcal{R}(\boldsymbol{z})\} \,|\, \boldsymbol{z}, \mu, \tau) = \prod_{i,\; \boldsymbol{s}_i \in \mathcal{R}(\boldsymbol{z})} \mathcal{N}_{\mathrm{t}}(\theta_i\,;\, \mu, \tau).$$

Note: τ is a measure of defect-signal variability;
for example, if $\tau = 0 \implies$ all θ_i within the defect region are *equal to* μ.



Defect-Signal (Reflectivity) Model: An Illustration.



Defect-Signal (Reflectivity) Model: An Illustration.

Location, Shape, and Defect-Signal Distribution Parameters

The vector of the location, shape, and defect-signal distribution parameters **(model parameters)** is

$$\boldsymbol{\phi} = \begin{bmatrix} \boldsymbol{z} \\ \mu \\ \tau \end{bmatrix}$$

where

- $\boldsymbol{z} = [x_{0,1}, x_{0,2}, d, A, \varphi]^T \equiv$ location and shape parameters and
- $\mu, \tau \equiv$ defect-signal distribution parameters.

Outline

- Measurement Model
  - Parametric model for defect location and shape,
  - Measurement-error (noise) model,
  - Defect-signal (reflectivity) model,
  - Prior specifications for the location, shape, and defect-signal distribution parameters (model parameters).
- Bayesian Analysis
  - Simulating the model parameters ϕ,
  - Simulating the signals θ_i.
- Numerical Examples.

Prior Specifications for the Model Parameters

$$\pi_{\boldsymbol{\phi}}(\boldsymbol{\phi}) = \pi_{x_{0,1}}(x_{0,1}) \cdot \pi_{x_{0,2}}(x_{0,2}) \cdot \pi_d(d) \cdot \pi_A(A) \cdot \pi_\varphi(\varphi) \cdot \pi_\mu(\mu) \cdot \pi_\tau(\tau)$$

where

$$\begin{aligned}
\pi_{x_{0,1}}(x_{0,1}) &= \text{uniform}(x_{0,1,\text{MIN}}, x_{0,1,\text{MAX}}) \\
\pi_{x_{0,2}}(x_{0,2}) &= \text{uniform}(x_{0,2,\text{MIN}}, x_{0,2,\text{MAX}}) \\
\pi_d(d) &= \text{uniform}(d_{\text{MIN}}, d_{\text{MAX}}) \\
\pi_A(A) &= \text{uniform}(A_{\text{MIN}}, A_{\text{MAX}}), \quad \pi_\varphi(\varphi) = \text{uniform}(\varphi_{\text{MIN}}, \varphi_{\text{MAX}}) \\
\pi_\mu(\mu) &= \text{uniform}(0, \mu_{\text{MAX}}), \quad \pi_\tau(\tau) = \text{uniform}(0, \tau_{\text{MAX}}).
\end{aligned}$$

Outline

- Measurement Model
  - Parametric model for defect location and shape,
  - Measurement-error (noise) model,
  - Defect-signal (reflectivity) model,
  - Prior specifications for the location, shape, and defect-signal distribution parameters (model parameters).

- Bayesian Analysis
  - Simulating the model parameters ϕ,
  - Simulating the signals θ_i.

- Numerical Examples.

Simulating the Model Parameters ϕ

Sample from the posterior pdf

$$p(\boldsymbol{\phi} \,|\, \boldsymbol{y})$$

where $\boldsymbol{y} = [y_1, y_2, \ldots, y_{\mathbb{N}_{\mathrm{tot}}}]^T$ denotes the vector of all observations.

Comments: Obtaining a closed-form expression for $p(\boldsymbol{\phi} \,|\, \boldsymbol{y})$ is impossible, but we can determine it up to a multiplicative constant $\Longrightarrow$ sufficient for applying MCMC techniques and simulating $\boldsymbol{\phi}$s from $p(\boldsymbol{\phi} \,|\, \boldsymbol{y})$!

Finding $p(\boldsymbol{\phi} \,|\, \boldsymbol{y})$ (up to a Multiplicative Constant)

Define the vector of random signals $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_{\mathbb{N}_{\mathrm{tot}}}]^T$.

Idea. *Integrate out the random signals $\boldsymbol{\theta}$* from the joint posterior pdf $p(\boldsymbol{\phi}, \boldsymbol{\theta} \,|\, \boldsymbol{y})$:

$$p(\boldsymbol{\phi} \,|\, \boldsymbol{y}) = \int p(\boldsymbol{\phi}, \boldsymbol{\theta} \,|\, \boldsymbol{y}) \, d\boldsymbol{\theta} = \frac{p(\boldsymbol{\phi}, \boldsymbol{\theta} \,|\, \boldsymbol{y})}{p(\boldsymbol{\theta} \,|\, \boldsymbol{\phi}, \boldsymbol{y})}.$$

Finding $p(\boldsymbol{\phi} \,|\, \boldsymbol{y})$ (up to a Multiplicative Constant)

Define the vector of random signals $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_{\mathbb{N}_{\mathrm{tot}}}]^T$.

Idea. *Integrate out the random signals* $\boldsymbol{\theta}$ from the joint posterior pdf $p(\boldsymbol{\phi}, \, \boldsymbol{\theta} \,|\, \boldsymbol{y})$:

$$p(\boldsymbol{\phi} \,|\, \boldsymbol{y}) = \int p(\boldsymbol{\phi}, \, \boldsymbol{\theta} \,|\, \boldsymbol{y}) \, d\boldsymbol{\theta} = \frac{p(\boldsymbol{\phi}, \, \boldsymbol{\theta} \,|\, \boldsymbol{y})}{p(\boldsymbol{\theta} \,|\, \boldsymbol{\phi}, \, \boldsymbol{y})}.$$

Note: A Bayesian trick for "integrating" $\boldsymbol{\theta}$ out without actually performing the integration!

Computing $p(\boldsymbol{\phi}, \boldsymbol{\theta} \mid \boldsymbol{y})$

$$
\begin{aligned}
p(\boldsymbol{\phi},\, \boldsymbol{\theta} \mid \boldsymbol{y}) \quad &\propto \quad \pi_{\boldsymbol{\phi}}(\boldsymbol{\phi}) \cdot p(\boldsymbol{\theta} \mid \boldsymbol{\phi}) \cdot p(\boldsymbol{y} \mid \boldsymbol{\theta}) \\
&\propto \quad \pi_{\boldsymbol{\phi}}(\boldsymbol{\phi}) \cdot \prod_{i,\ \boldsymbol{s}_i \in \mathcal{R}(\boldsymbol{z})} \mathcal{N}_{\mathrm{t}}(\theta_i; \mu, \tau) \\
&\qquad \cdot \prod_{i,\, \boldsymbol{s}_i \in \mathcal{R}(\boldsymbol{z})} \mathcal{N}(y_i; \theta_i, \sigma) \cdot \prod_{j,\, \boldsymbol{s}_j \in \mathcal{R}^{\mathrm{c}}(\boldsymbol{z})} \mathcal{N}(y_j; 0, \sigma) \\
&\propto \quad \pi_{\boldsymbol{\phi}}(\boldsymbol{\phi}) \cdot \Big[\prod_{i,\ \boldsymbol{s}_i \in \mathcal{R}(\boldsymbol{z})} \mathcal{N}_{\mathrm{t}}(\theta_i; \mu, \tau) \cdot \frac{\mathcal{N}(y_i; \theta_i, \sigma)}{\mathcal{N}(y_i; 0, \sigma)} \Big].
\end{aligned}
$$

Computing $p(\boldsymbol{\theta} \,|\, \boldsymbol{\phi}, \boldsymbol{y})$

$$
\begin{aligned}
p(\boldsymbol{\theta} \,|\, \boldsymbol{\phi},\, \boldsymbol{y}) \quad &\propto \quad \prod_{i,\ \boldsymbol{s}_i \in \mathcal{R}(\boldsymbol{z})} \exp\left[-\frac{(\theta_i - \mu)^2}{2\tau^2} - \frac{(y_i - \theta_i)^2}{2\sigma^2} \right] \cdot i_{[0,\infty)}(\theta_i) \\
&= \quad \prod_{i,\ \boldsymbol{s}_i \in \mathcal{R}(\boldsymbol{z})} \mathcal{N}_{\mathrm{t}}\left(\theta_i \,;\, \widetilde{\theta}_i(\mu, \tau), \left(\frac{1}{\tau^2} + \frac{1}{\sigma^2} \right)^{-1/2} \right)
\end{aligned}
$$

where

$$
\widetilde{\theta}_i(\mu, \tau) \quad = \quad \frac{\tau^2\, y_i + \sigma^2\, \mu}{\tau^2 + \sigma^2}.
$$

Finally $p(\boldsymbol{\phi} \,|\, \boldsymbol{y})$!

$$
\begin{aligned}
p(\boldsymbol{\phi} \,|\, \boldsymbol{y}) \quad &= \quad \frac{p(\boldsymbol{\phi},\, \boldsymbol{\theta} \,|\, \boldsymbol{y})}{p(\boldsymbol{\theta} \,|\, \boldsymbol{\phi},\, \boldsymbol{y})} \\
&\propto \quad \pi_{\boldsymbol{\phi}}(\boldsymbol{\phi}) \cdot \overbrace{\prod_{i,\ \boldsymbol{s}_i \in \mathcal{R}(\boldsymbol{z})} \frac{\mathcal{N}_{\mathrm{t}}(\theta_i; \mu, \tau)}{\mathcal{N}_{\mathrm{t}}\Big(\theta_i;\, \widetilde{\theta}_i(\mu, \tau), (1/\tau^2 + 1/\sigma^2)^{-1/2}\Big)} \cdot \frac{\mathcal{N}(y_i; \theta_i, \sigma)}{\mathcal{N}(y_i; 0, \sigma)}}^{l(\boldsymbol{y} \,|\, \boldsymbol{\phi})}.
\end{aligned}
$$

Here, $l(\boldsymbol{y} \,|\, \boldsymbol{\phi}) \equiv$ *likelihood function* of $\boldsymbol{\phi}$.

Simplifying $l(\boldsymbol{y} \mid \boldsymbol{\phi})$

$l(\boldsymbol{y} \mid \boldsymbol{\phi})$ *must not* depend on $\boldsymbol{\theta} \Longrightarrow$ (arbitrarily) set $\theta_i = \mu$ in $l(\boldsymbol{y} \mid \boldsymbol{\phi})$ on the previous page, yielding

$$l(\boldsymbol{y} \mid \boldsymbol{\phi}) \quad = \quad \prod_{i,\ \boldsymbol{s}_i \in \mathcal{R}(\boldsymbol{z})} \frac{\mathcal{N}_{\mathrm{t}}(\mu\,;\,\mu, \tau)}{\mathcal{N}_{\mathrm{t}}\Big(\mu\,;\,\widetilde{\theta}_i(\mu, \tau), (1/\tau^2 + 1/\sigma^2)^{-1/2}\Big)} \cdot \frac{\mathcal{N}(y_i; \mu, \sigma)}{\mathcal{N}(y_i; 0, \sigma)}.$$

Log Likelihood

$$\ln l(\boldsymbol{y}\,|\,\boldsymbol{\phi}) \;=\; -\frac{N(\boldsymbol{z})}{2}\,\ln\left(1+\frac{\tau^2}{\sigma^2}\right) \\ \qquad + \sum_{i,\;\boldsymbol{s}_i\in\mathcal{R}(\boldsymbol{z})}\left\{\ln\left[\frac{\Phi\left(\widetilde{\theta}_i(\mu,\tau)\cdot\sqrt{\frac{1}{\tau^2}+\frac{1}{\sigma^2}}\right)}{\Phi(\mu/\tau)}\right]+\frac{y_i^2}{2\,\sigma^2}-\frac{(y_i-\mu)^2}{2\,(\tau^2+\sigma^2)}\right\}.$$

where

$N(\boldsymbol{z}) = \sum_{i,\,\boldsymbol{s}_i\in\mathcal{R}(\boldsymbol{z})} 1 \quad \equiv \quad$ number of measurements collected over $\mathcal{R}(\boldsymbol{z})$.

Log Likelihood: Comments

$$\ln l(\boldsymbol{y} \,|\, \boldsymbol{\phi}) \;=\; -\frac{N(\boldsymbol{z})}{2} \ln\Big(1 + \frac{\tau^2}{\sigma^2}\Big) \\ + \sum_{i,\ \boldsymbol{s}_i \in \mathcal{R}(\boldsymbol{z})} \Bigg\{ \ln\Bigg[\frac{\Phi\Big(\widetilde{\theta}_i(\mu,\tau)\cdot\sqrt{\frac{1}{\tau^2}+\frac{1}{\sigma^2}}\Big)}{\Phi(\mu/\tau)}\Bigg] + \frac{y_i^2}{2\,\sigma^2} - \frac{(y_i-\mu)^2}{2\,(\tau^2+\sigma^2)} \Bigg\}.$$

If we set $\ln\Big[\frac{\Phi\big(\widetilde{\theta}_i(\mu,\tau)\cdot\sqrt{\frac{1}{\tau^2}+\frac{1}{\sigma^2}}\big)}{\Phi(\mu/\tau)}\Big]$ to zero in the above expression (i.e. allow θ_i to be negative), we obtain the normalized log likelihood in [1] and our approach reduces to that in [1].

[1] A. Dogandžić and B. Zhang, "Bayesian NDE defect signal analysis," *IEEE Trans. Signal Processing*, vol. 55, pp. 372–378, Jan. 2007.

Outline

- Measurement Model
  - Parametric model for defect location and shape,
  - Measurement-error (noise) model,
  - Defect-signal (reflectivity) model,
  - Prior specifications for the location, shape, and defect-signal distribution parameters (model parameters).
- Bayesian Analysis
  - Simulating the model parameters $\boldsymbol{\phi}$,
  - Simulating the signals θ_i.
- Numerical Examples.

Background: Gibbs Sampler

For simplicity, consider sampling scalar random variables ϕ and u from their joint pdf $p(\phi, u)$. We may know this joint pdf up to a multiplicative constant only, say $p(\phi, u) \propto h(\phi, u)$. Suppose that

- we cannot sample directly from $p(\phi, u) \propto h(\phi, u)$ but
- we can sample from the conditional pdfs

$$p({\color{red}\phi} \,|\, u) \propto h({\color{red}\phi}, u) \quad \text{and} \quad p({\color{blue}u} \,|\, \phi) \propto h(\phi, {\color{blue}u}).$$

Note: In many practical applications, $p({\color{red}\phi} \,|\, u)$ and $p({\color{blue}u} \,|\, \phi)$ are "standard" (and hence easy to sample from) or, if nonstandard, can be simulated using von Neumann's rejection method.

Gibbs Sampler

Do the following:

(a) Start at some $\phi^{(0)}$ and $u^{(0)}$;

(b) For $t = 1, 2, \ldots$

Step 1: Sample $u^{(t)}$ from $h(\phi^{(t-1)}, \cdot)$ and
Step 2: Sample $\phi^{(t)}$ from $h(\cdot, u^{(t)})$

[i.e. create $(\phi^{(t)}, u^{(t)})$ using $(\phi^{(t-1)}, u^{(t-1)})$ $\Longrightarrow$ Markov chain!].

Under appropriate circumstances, $(\phi^{(1)}, u^{(1)}), (\phi^{(2)}, u^{(2)}), \ldots, (\phi^{(T)}, u^{(T)})$ can be used to approximate properties of $p(\phi, u) \propto h(\phi, u)$!

Gibbs Sampler

Do the following:

(a) Start at some $\phi^{(0)}$ and $u^{(0)}$;

(b) For $t = 1, 2, \ldots$

Step 1: Sample $u^{(t)}$ from $h(\phi^{(t-1)}, \cdot)$ and
Step 2: Sample $\phi^{(t)}$ from $h(\cdot, u^{(t)})$

[i.e. create $(\phi^{(t)}, u^{(t)})$ using $(\phi^{(t-1)}, u^{(t-1)}) \Longrightarrow$ Markov chain!].

Under appropriate circumstances, $(\phi^{(1)}, u^{(1)}), (\phi^{(2)}, u^{(2)}), \ldots, (\phi^{(T)}, u^{(T)})$ can be used to approximate properties of $p(\phi, u) \propto h(\phi, u)$!

The draws $(\phi^{(t)}, u^{(t)})$ are not i.i.d. but we do not care (much)!

Gibbs Sampler: Comment

Gibbs sampler is named after J.W. Gibbs, a 19th century American physicist and mathematician and one of the founders of modern thermodynamics and statistical mechanics.

Gibbs Sampler: Comment

Gibbs sampler is named after J.W. Gibbs, a 19th century American physicist and mathematician and one of the founders of modern thermodynamics and statistical mechanics.

But, Gibbs did not invent the Gibbs sampler. A more descriptive name has been proposed: *successive substitution sampling*. Yet, the name "Gibbs sampler" has won.

Gibbs Sampler: Comment (cont.)

The "Gibbs sampler" is yet another example of Stigler's Law of Eponymy, which states that

No scientific discovery is named after the person(s) who thought of it.

Gibbs Sampler: Comment (cont.)

The "Gibbs sampler" is yet another example of Stigler's Law of Eponymy, which states that

No scientific discovery is named after the person(s) who thought of it.

Interestingly, Stigler's Law of Eponymy is not due to Stigler [2], meaning that it is an example of itself!

[2] S. Stigler, *Statistics on the Table: The History of Statistical Concepts and Methods*, Cambridge, MA: Harvard University Press, 1999.

Background: (Univariate) Slice Sampler

Consider now sampling a random variable ϕ from a nonstandard $p(\phi) \propto h(\phi)$.

(Seemingly Counter-Intuitive!) Idea:

- Invent a convenient bivariate distribution for, say, ϕ and u, with marginal pdf for ϕ specified by $h(\phi)$.

- Then, use Gibbs sampling to make

$$(\phi^{(0)}, u^{(0)}), (\phi^{(1)}, u^{(1)}), (\phi^{(2)}, u^{(2)}), \ldots, (\phi^{(T)}, u^{(T)}).$$

Background: (Univariate) Slice Sampler

Consider now sampling a random variable ϕ from a nonstandard $p(\phi) \propto h(\phi)$.

(Seemingly Counter-Intuitive!) Idea:

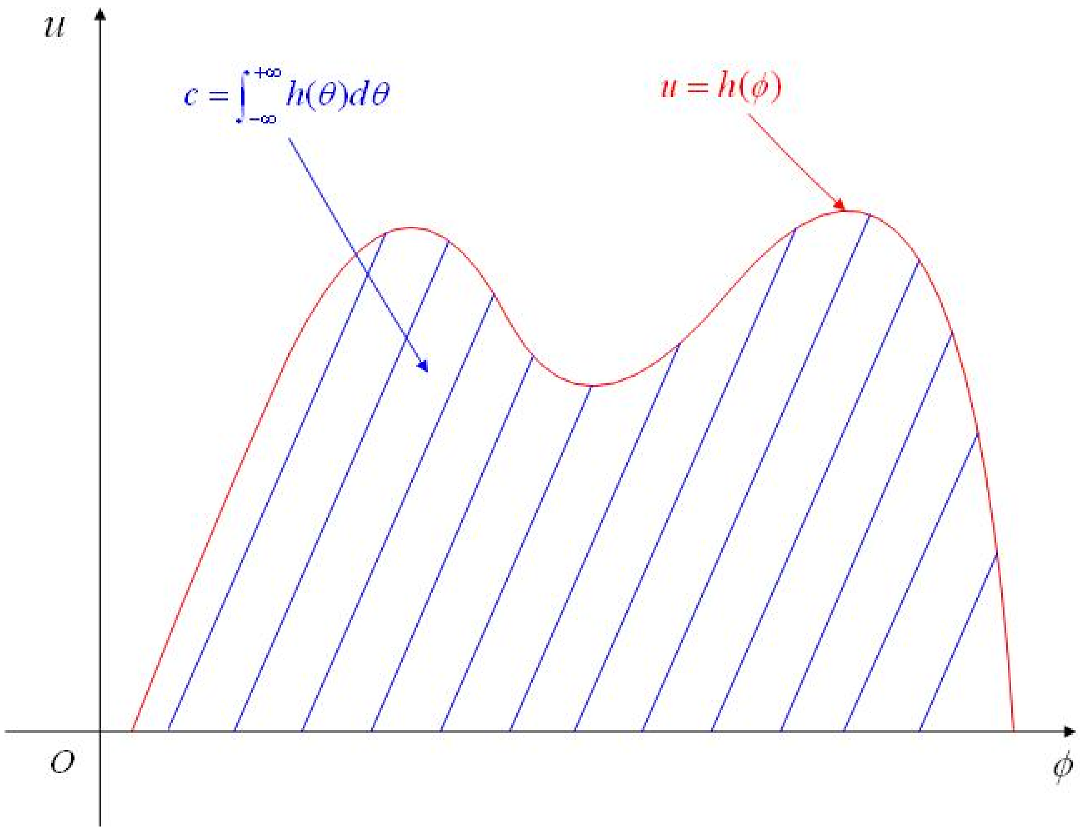
- Invent a convenient bivariate distribution for, say, ϕ and u, with marginal pdf for ϕ specified by $h(\phi)$.

- Then, use Gibbs sampling to make

$$(\phi^{(0)}, u^{(0)}), (\phi^{(1)}, u^{(1)}), (\phi^{(2)}, u^{(2)}), \ldots, (\phi^{(T)}, u^{(T)}).$$

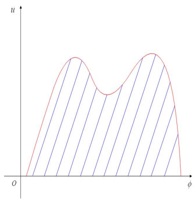
Create an auxiliary variable u just for convenience!

(Univariate) Slice Sampler



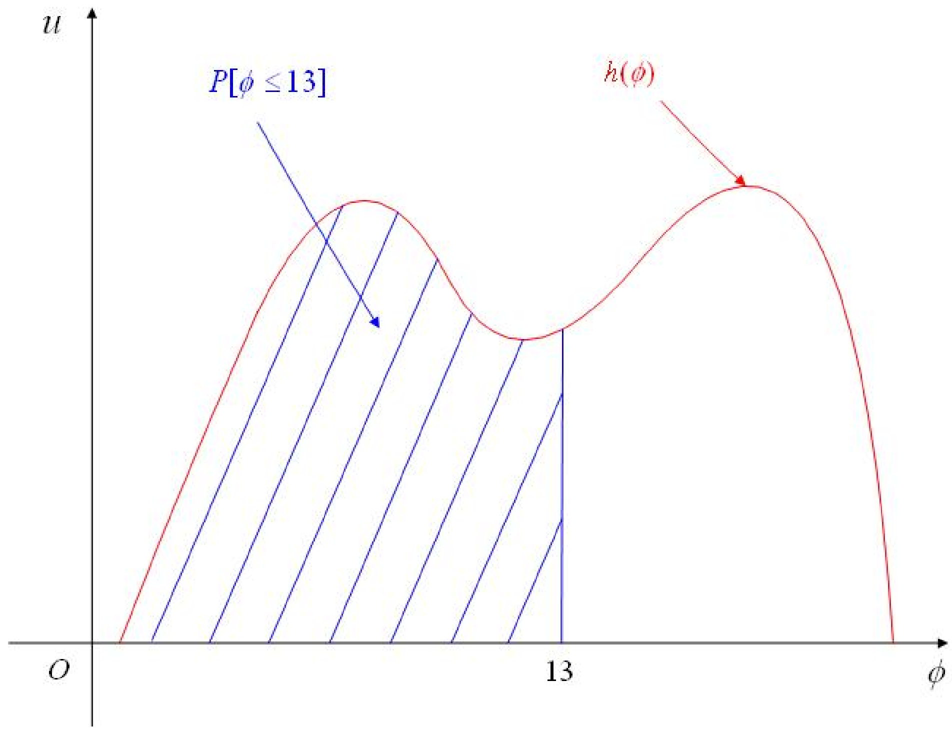
(Univariate) Slice Sampler

"Invent" a joint distribution for ϕ and u by declaring it to be uniform on

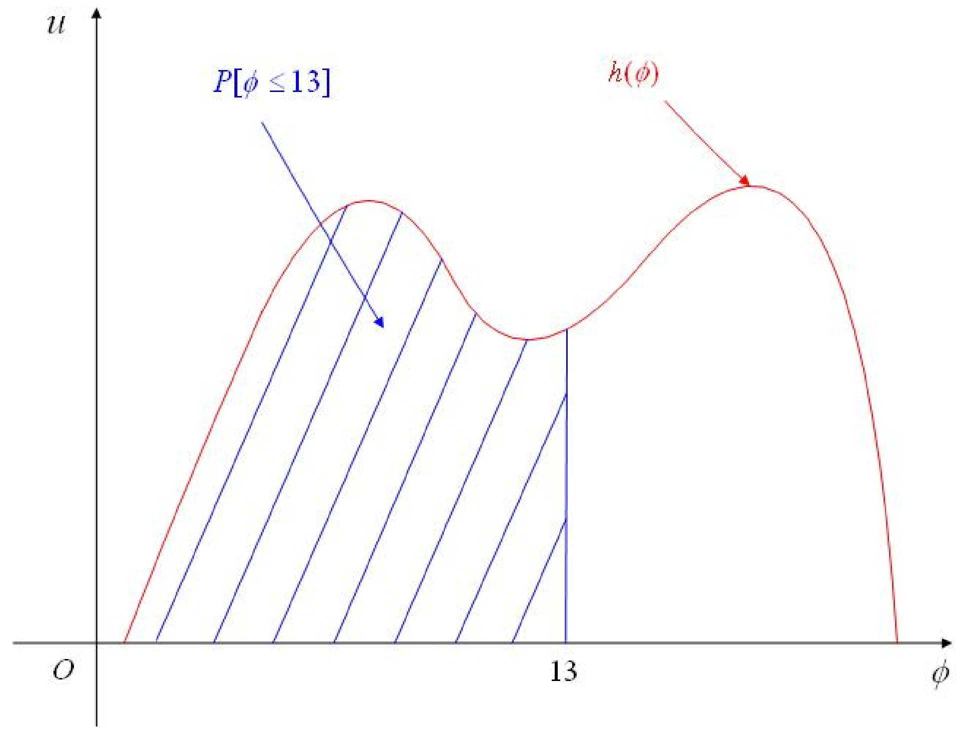


:

$$p(\phi, u) = \begin{cases} \frac{1}{c}, & 0 < u < h(\phi) \\ 0, & \text{otherwise} \end{cases} \;\propto\; i_{(0,h(\phi))}(u).$$



With this joint pdf, $P[\phi \leq 13] = \int_{-\infty}^{13} \frac{h(\phi)}{c} \, d\phi$.



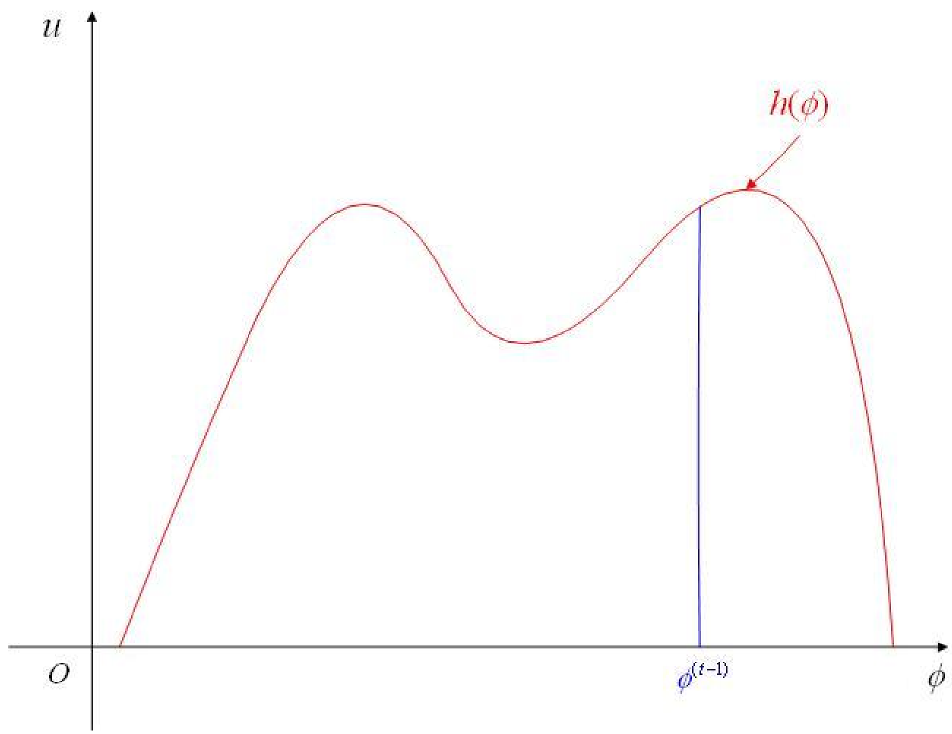
With this joint pdf, $P[\phi \leq 13] = \int_{-\infty}^{13} \frac{h(\phi)}{c}\, d\phi$.

The marginal pdf of ϕ is indeed specified by $h(\phi) \Longrightarrow$ if we figure out how to do Gibbs sampling, we know how to generate a ϕ from $h(\phi)$.

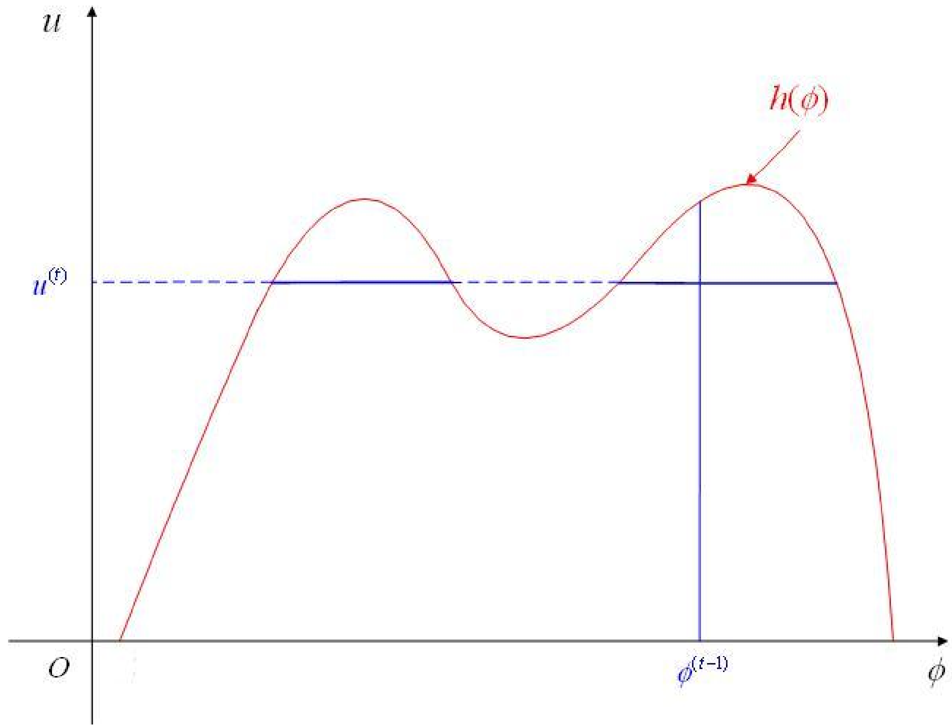
Gibbs Sampler is Easy in This Case!

$$
\begin{aligned}
p(u \,|\, \phi) &= \text{uniform}\Big(0, h(\phi)\Big) \\
p(\phi \,|\, u) &= \text{uniform on } \underbrace{\{\phi \,|\, h(\phi) > u\}}_{\text{“slice”}}.
\end{aligned}
$$

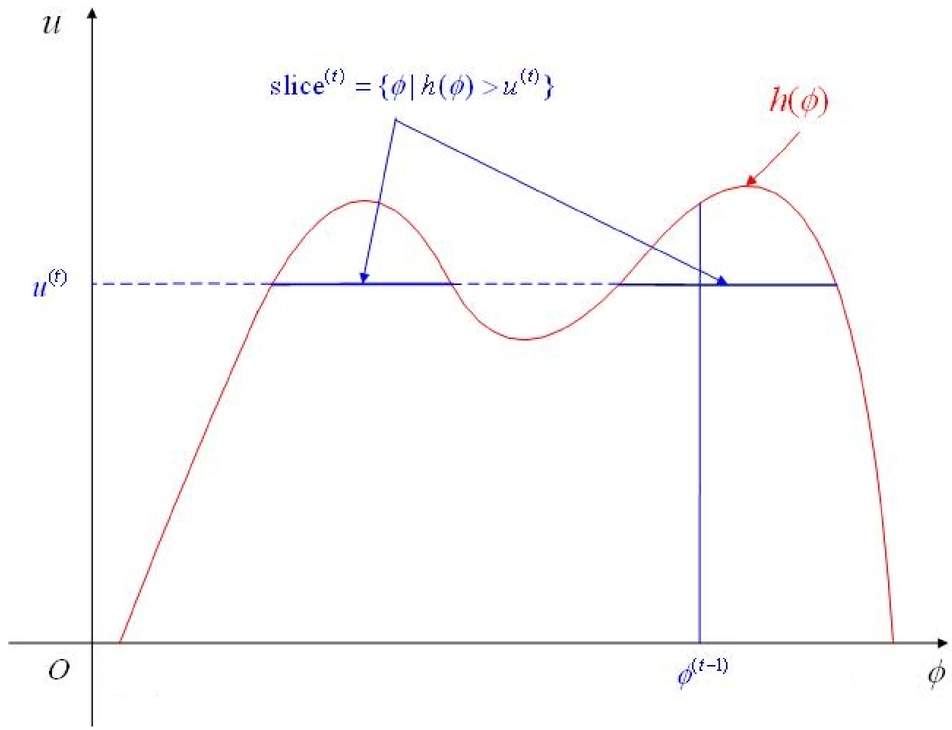
Step 1: Given $\phi^{(t-1)}$, Sample $u^{(t)} \sim \mathrm{uniform}\Big(0, h(\phi^{(t-1)})\Big)$



Step 1: Given $\phi^{(t-1)}$, Sample $u^{(t)} \sim \mathrm{uniform}\Big(0, h(\phi^{(t-1)})\Big)$



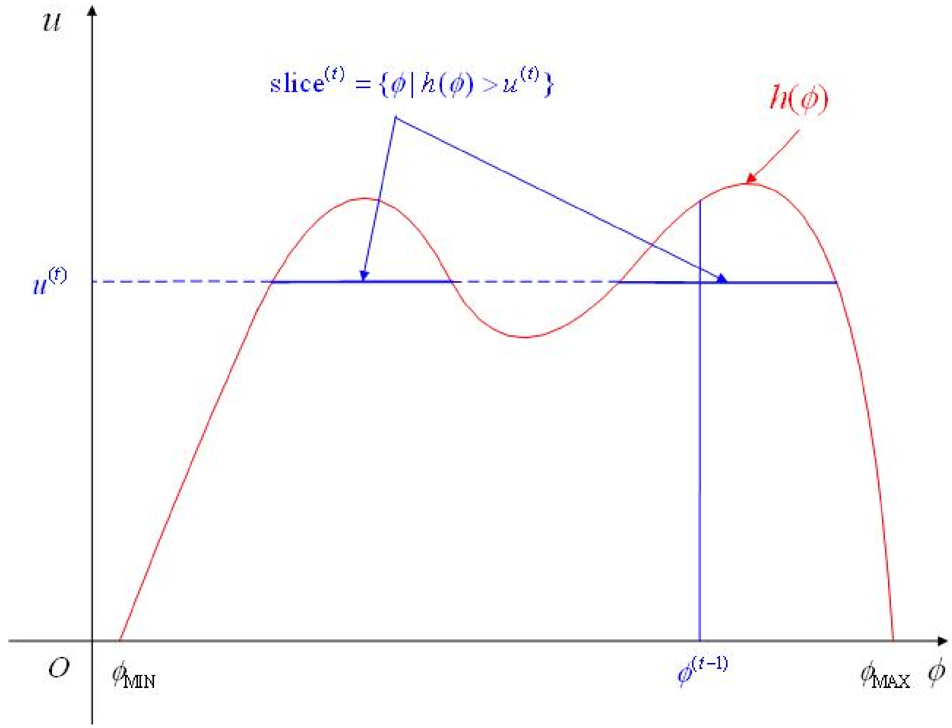
Step 2: Given $u^{(t)}$, Sample $\phi^{(t)}$ Uniform from slice$^{(t)}$



If we can algebraically solve $h(\phi) = u^{(t)}$, our task is easy. What if not?

Step 2 Implementation Using the Rejection Method

When we have band bounds on ϕ, say $\phi_{\mathrm{MIN}} \leq \phi \leq \phi_{\mathrm{MAX}}$



generate i.i.d. values ϕ from uniform$(\phi_{\mathrm{MIN}}, \phi_{\mathrm{MAX}})$ until we produce a ϕ in the slice [i.e. $h(\phi) > u^{(t)}$], which we then accept as $\phi^{(t)}$.

Back to Our Problem: Slice Sampling from $p(\phi \,|\, \boldsymbol{y}) \propto \boldsymbol{\pi}_{\phi}(\phi)\, \boldsymbol{l}(\boldsymbol{y} \,|\, \phi)$

Create an auxiliary variable u so that

$$p(\boldsymbol{\phi}, u \,|\, \boldsymbol{y}) \propto \pi_{\boldsymbol{\phi}}(\boldsymbol{\phi}) \cdot i_{(0,\, l(\boldsymbol{y} \,|\, \boldsymbol{\phi}))}(u)$$

implying

$$\begin{aligned} p(u \,|\, \boldsymbol{\phi}, \boldsymbol{y}) &= \mathrm{uniform}\Big(0,\, l(\boldsymbol{y} \,|\, \boldsymbol{\phi})\Big) \\ p(\boldsymbol{\phi} \,|\, u, \boldsymbol{y}) &\propto \pi_{\boldsymbol{\phi}}(\boldsymbol{\phi}) \cdot i_{(0,\, l(\boldsymbol{y} \,|\, \boldsymbol{\phi}))}(u). \end{aligned}$$

Slice Sampling from $\boldsymbol{p}(\boldsymbol{\phi} \,|\, \boldsymbol{y}) \propto \boldsymbol{\pi}_{\phi}(\boldsymbol{\phi})\, \boldsymbol{l}(\boldsymbol{y} \,|\, \boldsymbol{\phi})$

We can easily sample from $p(\boldsymbol{\phi}\,, u|\, \boldsymbol{y})$ using Gibbs, whose one cycle is given by the following two steps:

Step 1: Draw a $u^{(t)}$ from $\mathrm{uniform}\big(0, l(\boldsymbol{y} \,|\, \boldsymbol{\phi}^{(t-1)})\big)$ and

Step 2: Draw a $\boldsymbol{\phi}^{(t)}$ from its prior pdf $\pi_{\boldsymbol{\phi}}(\boldsymbol{\phi})$ *subject to* the indicator restriction $l(\boldsymbol{y} \,|\, \boldsymbol{\phi}^{(t)}) \geq u^{(t)}$.

Comments

- If a vector ϕ satisfies the indicator restriction:

$$l(\boldsymbol{y} \mid \boldsymbol{\phi}) \geq u^{(t)}, \qquad \text{we say that it is } \textit{in the slice}.$$

- Step 2 is called "getting a point in the slice," corresponding to sampling from

$$p(\boldsymbol{\phi} \mid u^{(t)}, \boldsymbol{y}) \propto \pi_{\boldsymbol{\phi}}(\boldsymbol{\phi}) \cdot i_{(0,\, l(\boldsymbol{y} \mid \boldsymbol{\phi}))}(u^{(t)}).$$

- A "naive" rejection method for getting a point in the slice: keep drawing ϕs i.i.d. from $\pi_{\phi}(\phi)$ *until* we get a ϕ that is *in the slice*. This may take forever since ϕ is a 7-D vector!

Shrinkage Sampling for Getting a Point in the Slice [3]

Recall: The parameter space of $\boldsymbol{\phi} = [x_{0,1}, x_{0,2}, d, A, \varphi, \mu, \tau]^T$ is a *hyperrectangle* with $x_{0,1} \in (x_{0,1,\mathrm{MIN}}, x_{0,1,\mathrm{MAX}})$, $x_{0,2} \in (x_{0,2,\mathrm{MIN}}, x_{0,2,\mathrm{MAX}})$, $d \in (d_{\mathrm{MIN}}, d_{\mathrm{MAX}})$, $A \in (A_{\mathrm{MIN}}, A_{\mathrm{MAX}})$, $\phi \in (-\pi/4, \pi/4)$, $\mu \in (0, \mu_{\mathrm{MAX}})$, and $\tau \in (0, \tau_{\mathrm{MAX}})$.

Shrinkage Sampling for Getting a Point in the Slice [3]

Recall: The parameter space of $\boldsymbol{\phi} = [x_{0,1}, x_{0,2}, d, A, \varphi, \mu, \tau]^T$ is a *hyperrectangle* with $x_{0,1} \in (x_{0,1,\mathrm{MIN}}, x_{0,1,\mathrm{MAX}})$, $x_{0,2} \in (x_{0,2,\mathrm{MIN}}, x_{0,2,\mathrm{MAX}})$, $d \in (d_{\mathrm{MIN}}, d_{\mathrm{MAX}})$, $A \in (A_{\mathrm{MIN}}, A_{\mathrm{MAX}})$, $\phi \in (-\pi/4, \pi/4)$, $\mu \in (0, \mu_{\mathrm{MAX}})$, and $\tau \in (0, \tau_{\mathrm{MAX}})$.

Shrink the hyperrectangle from which the $\boldsymbol{\phi}$s are sampled in such a way that the previous value $\boldsymbol{\phi}^{(t-1)}$ remains within the hyperrectangle [3].

Note: $\boldsymbol{\phi}^{(t-1)}$ is always in the slice since $u^{(t)} \sim \mathrm{uniform}\big(0, l(\boldsymbol{y} \,|\, \boldsymbol{\phi}^{(t-1)})\big)$!

---

[3] R.M. Neal, "Slice sampling," *Ann. Statist.*, vol. 31, pp. 705–741, June 2003.

Shrinkage Sampling from $p(\boldsymbol{\phi} \mid u^{(t)}, \boldsymbol{y})$

We first define the initial (largest) hyperrectangle with limits

$$
\begin{aligned}
x_{0,1,\mathrm{L}} &= x_{0,1,\mathrm{MIN}}, \quad x_{0,1,\mathrm{U}} = x_{0,1,\mathrm{MAX}}, \\
x_{0,2,\mathrm{L}} &= x_{0,2,\mathrm{MIN}}, \quad x_{0,2,\mathrm{U}} = x_{0,2,\mathrm{MAX}}, \\
d_{\mathrm{L}} &= d_{\mathrm{MIN}}, \quad d_{\mathrm{U}} = d_{\mathrm{MAX}} \\
A_{\mathrm{L}} &= A_{\mathrm{MIN}}, \quad A_{\mathrm{U}} = A_{\mathrm{MAX}} \\
\varphi_{\mathrm{L}} &= \varphi_{\mathrm{MIN}}, \quad \varphi_{\mathrm{U}} = \varphi_{\mathrm{MAX}} \\
\mu_{\mathrm{L}} &= 0, \quad \mu_{\mathrm{U}} = \mu_{\mathrm{MAX}} \\
\tau_{\mathrm{L}} &= 0, \quad \tau_{\mathrm{U}} = \tau_{\mathrm{MAX}}.
\end{aligned}
$$

Shrinkage Sampling from $p(\boldsymbol{\phi} \,|\, u^{(t)}, \boldsymbol{y})$ (cont.)

Obtain $\boldsymbol{\phi}^{(t)}$ as follows:

1. Sample

$$
\begin{aligned}
x_{0,1} &\quad \text{from} \quad \mathrm{uniform}(x_{0,1,\mathrm{L}}, x_{0,1,\mathrm{U}}) \\
x_{0,2} &\quad \text{from} \quad \mathrm{uniform}(x_{0,2,\mathrm{L}}, x_{0,2,\mathrm{U}}) \\
d &\quad \text{from} \quad \mathrm{uniform}(d_{\mathrm{L}}, d_{\mathrm{U}}) \\
A &\quad \text{from} \quad \mathrm{uniform}(A_{\mathrm{L}}, A_{\mathrm{U}}) \\
\varphi &\quad \text{from} \quad \mathrm{uniform}(\varphi_{\mathrm{L}}, \varphi_{\mathrm{U}}) \\
\mu &\quad \text{from} \quad \mathrm{uniform}(\mu_{\mathrm{L}}, \mu_{\mathrm{U}}) \\
\tau &\quad \text{from} \quad \mathrm{uniform}(\tau_{\mathrm{L}}, \tau_{\mathrm{U}})
\end{aligned}
$$

   yielding $\boldsymbol{\phi} = [x_{0,1}, x_{0,2}, d, A, \varphi, \mu, \tau]^T$.

2. Check if $\boldsymbol{\phi}$ is *within the slice*, i.e.

$$l(\boldsymbol{y} \mid \boldsymbol{\phi}) \geq u^{(t)}. \tag{1}$$

If (1) holds, return $\boldsymbol{\phi}^{(t)} = \boldsymbol{\phi}$ and exit the loop.

$\Big($Recall that $u^{(t)}$ was obtained by sampling from $\mathrm{uniform}\big(0, l(\boldsymbol{y} \mid \boldsymbol{\phi}^{(t-1)})\big)$.$\Big)$

3. If (1) does not hold, then *shrink the hyperrectangle*:

   - If $x_{0,1} < x_{0,1}^{(t-1)}$, set $x_{0,1,\mathrm{L}} = x_{0,1}$;
     else if $x_{0,1} > x_{0,1}^{(t-1)}$, set $x_{0,1,\mathrm{U}} = x_{0,1}$.
   - If $x_{0,2} < x_{0,2}^{(t-1)}$, set $x_{0,2,\mathrm{L}} = x_{0,2}$;
     else if $x_{0,2} > x_{0,2}^{(t-1)}$, set $x_{0,2,\mathrm{U}} = x_{0,2}$.
   - If $d < d^{(t-1)}$, set $d_{\mathrm{L}} = d$; else if $d > d^{(t-1)}$, set $d_{\mathrm{U}} = d$.
   - If $A < A^{(t-1)}$, set $A_{\mathrm{L}} = A$; else if $A > A^{(t-1)}$, set $A_{\mathrm{U}} = A$.
   - If $\varphi < \varphi^{(t-1)}$, set $\varphi_{\mathrm{L}} = \varphi$; else if $\varphi > \varphi^{(t-1)}$, set $\varphi_{\mathrm{U}} = \varphi$.
   - If $\mu < \mu^{(t-1)}$, set $\mu_{\mathrm{L}} = \mu$; else if $\mu > \mu^{(t-1)}$, set $\mu_{\mathrm{U}} = \mu$.
   - If $\tau < \tau^{(t-1)}$, set $\tau_{\mathrm{L}} = \tau$; else if $\tau > \tau^{(t-1)}$, set $\tau_{\mathrm{U}} = \tau$.
   - Go back to 1.

A Practical Modification of the Shrinkage Sampler

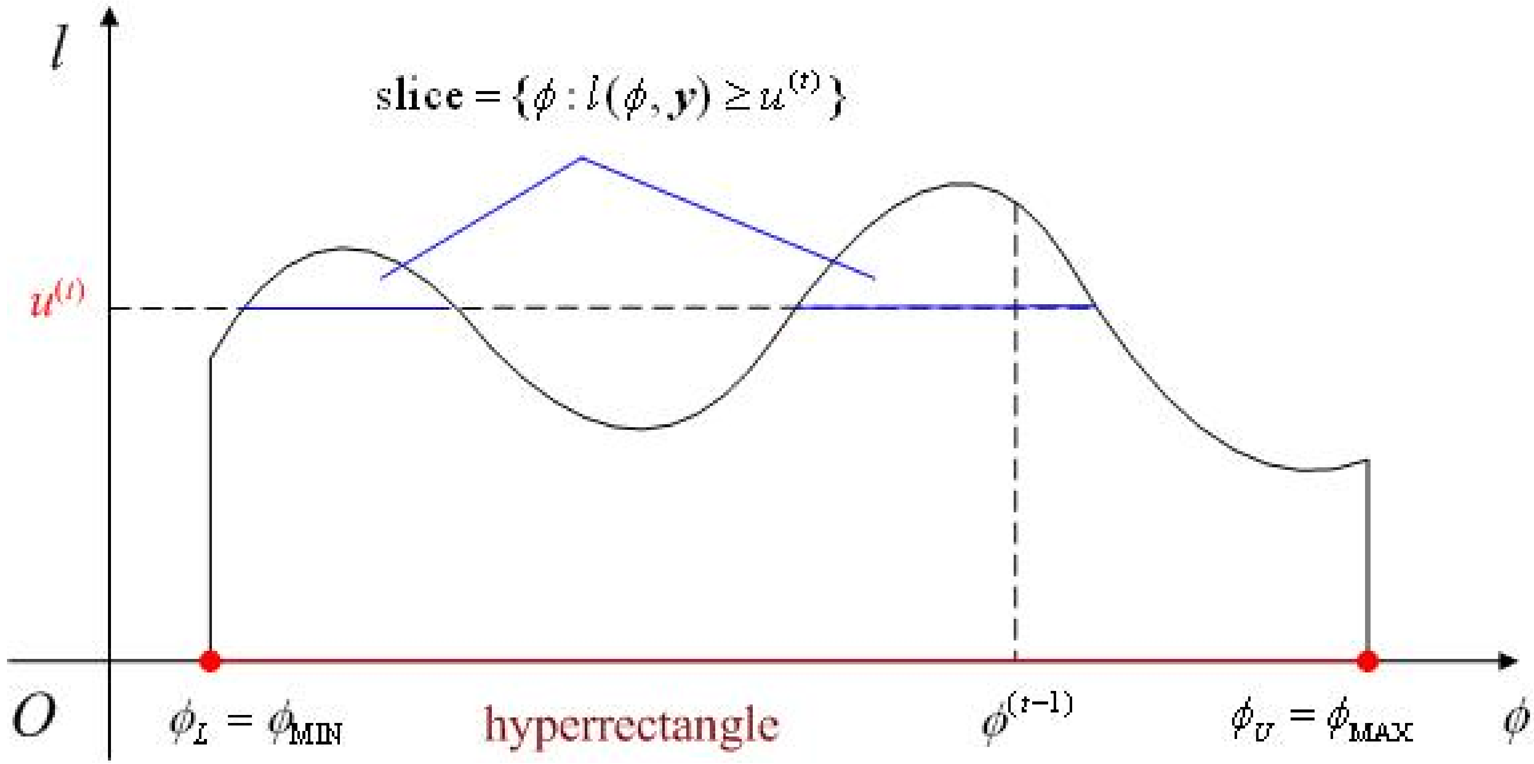
Computing $l(\boldsymbol{y} \,|\, \boldsymbol{\phi})$ may cause a floating-point underflow $\Longrightarrow$ safer to compute $\ln l(\boldsymbol{y} \,|\, \boldsymbol{\phi})$ rather than $l(\boldsymbol{y} \,|\, \boldsymbol{\phi})$. We then compute

$$\omega^{(t)} = \ln[l(\boldsymbol{y} \,|\, \boldsymbol{\phi}^{(t-1)})] - \epsilon$$

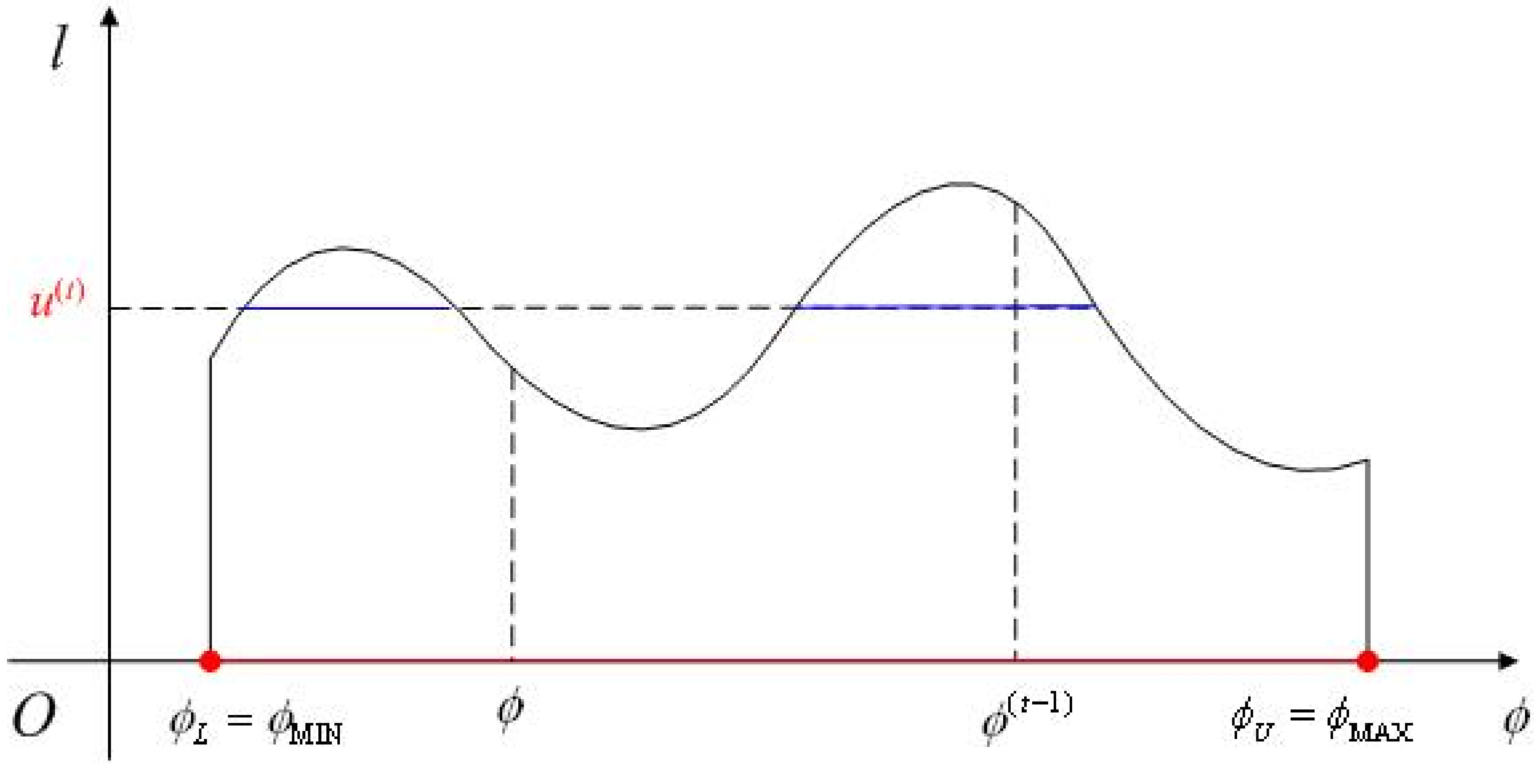
where ϵ is an exponential random variable with mean one. We say that $\boldsymbol{\phi}$ is *in the slice* if

$$\ln[l(\boldsymbol{y} \,|\, \boldsymbol{\phi})] \geq \omega^{(t)}$$

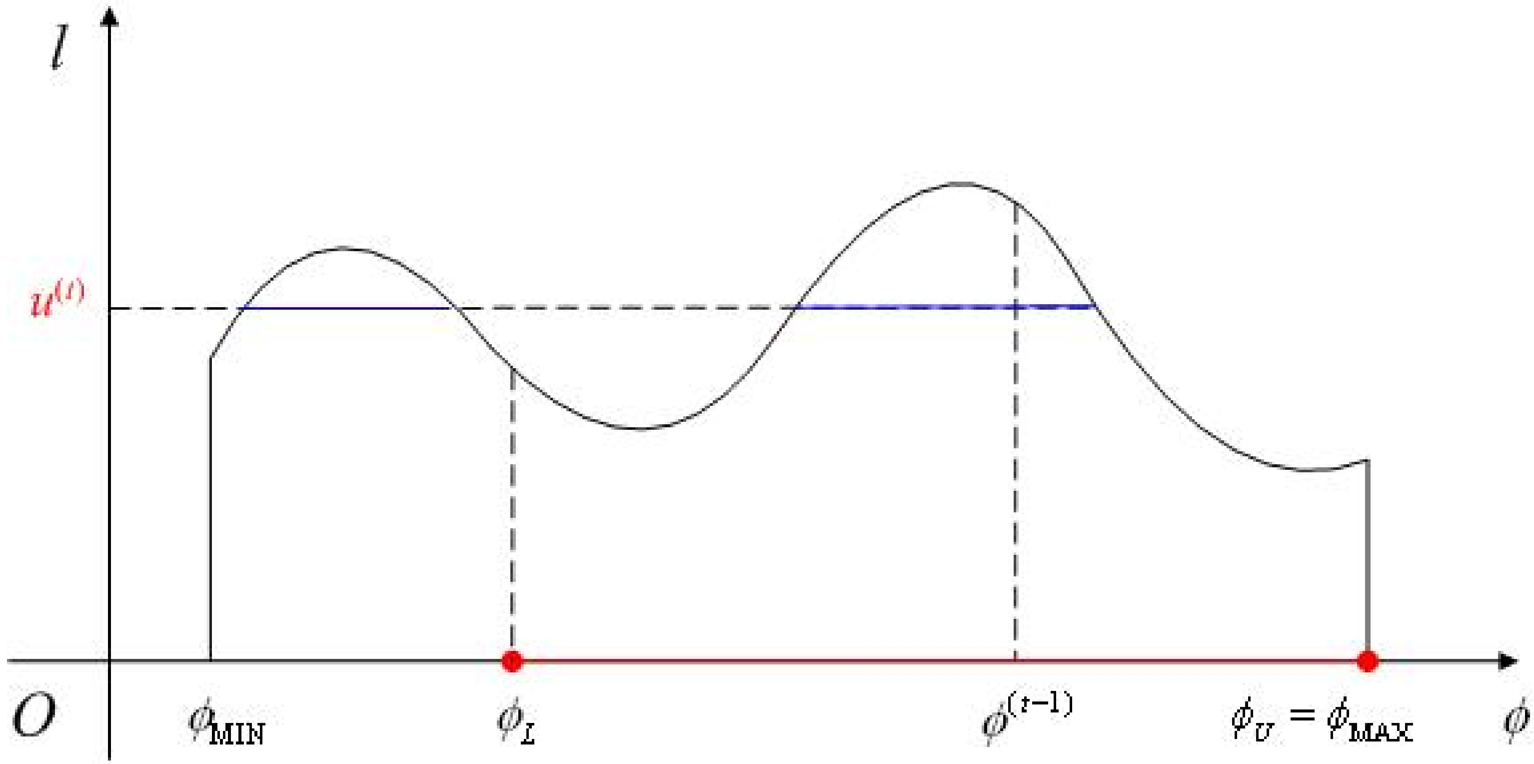
which is equivalent to (1).



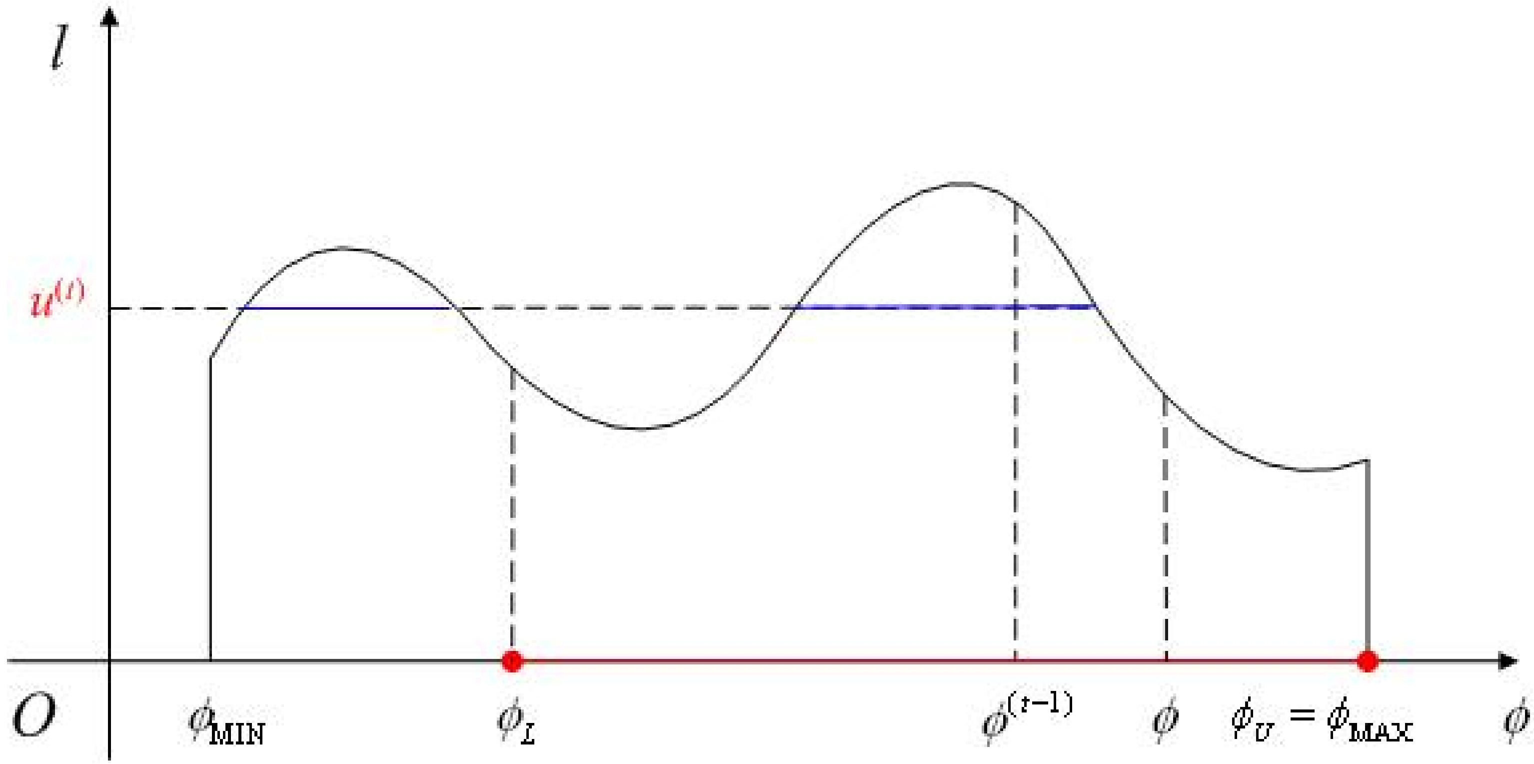
Shrinkage Slice Sampling: A 1-D Illustration.



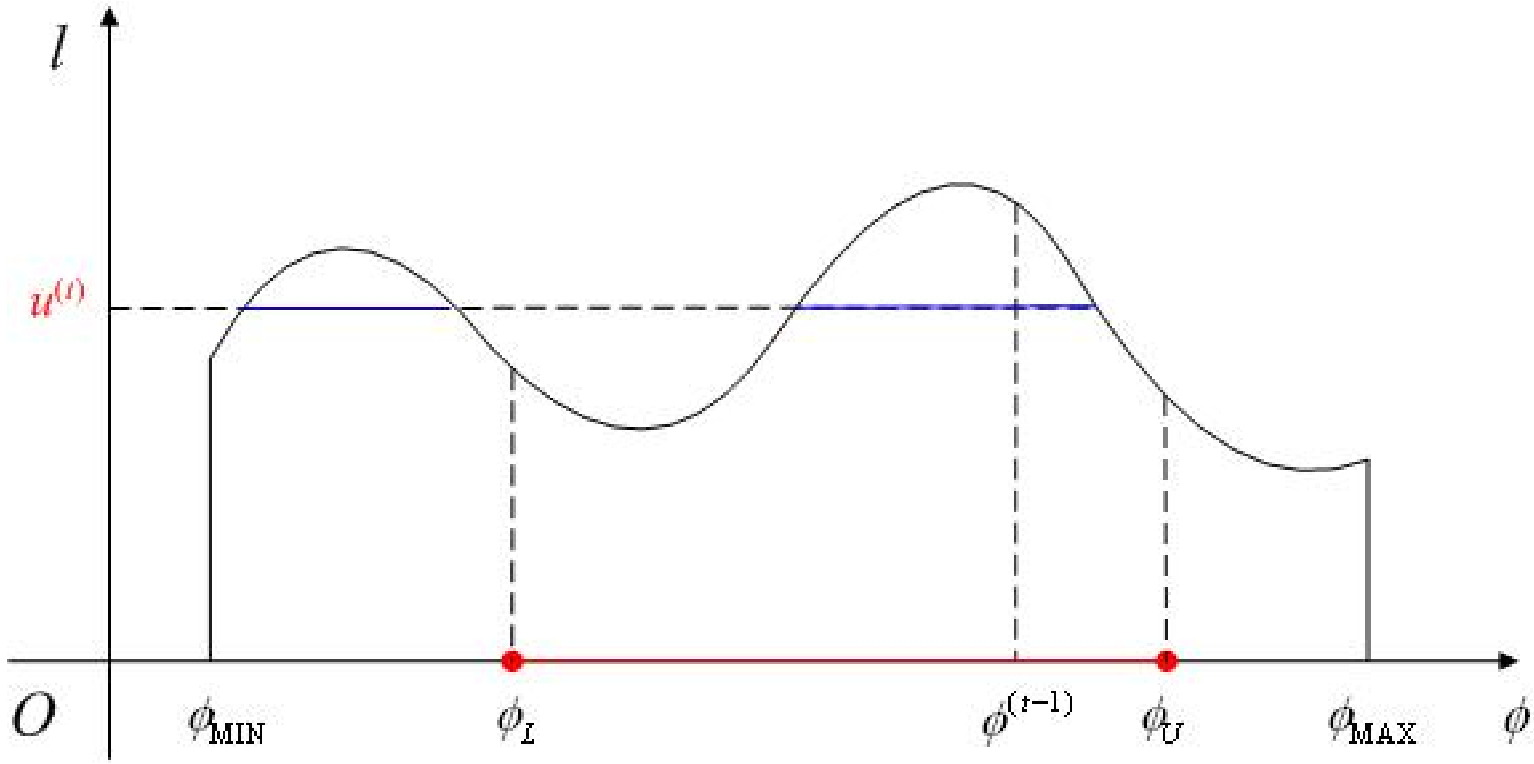
Shrinkage Slice Sampling: A 1-D Illustration.



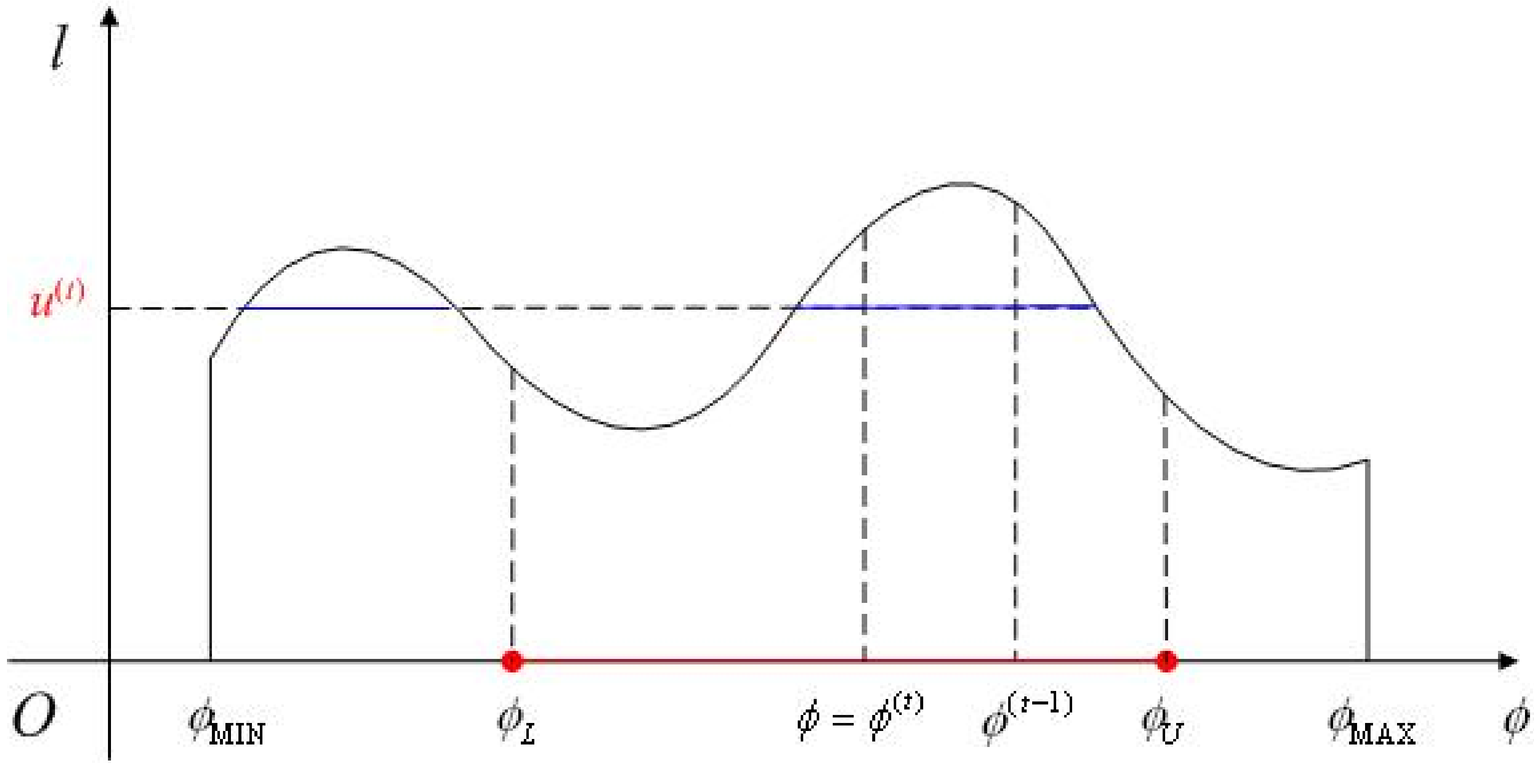
Shrinkage Slice Sampling: A 1-D Illustration.



Shrinkage Slice Sampling: A 1-D Illustration.



Shrinkage Slice Sampling: A 1-D Illustration.



Shrinkage Slice Sampling: A 1-D Illustration.

Outline

- Measurement Model
  - Parametric model for defect location and shape,
  - Measurement-error (noise) model,
  - Defect-signal (reflectivity) model,
  - Prior specifications for the location, shape, and defect-signal distribution parameters (model parameters).
- Bayesian Analysis
  - Simulating the model parameters ϕ,
  - Simulating the signals θ_i.
- Numerical Examples.

Simulating the Random Signals θ_i

To estimate the random signals $\boldsymbol{\theta}$, we utilize *composition sampling* from the posterior pdf $p(\boldsymbol{\theta} \,|\, \boldsymbol{y}) = \int p(\boldsymbol{\theta} \,|\, \boldsymbol{\phi}, \boldsymbol{y}) \, p(\boldsymbol{\phi} \,|\, \boldsymbol{y}) \, d\boldsymbol{\phi}$:

- Draw $\boldsymbol{\phi}^{(t)}$ from $p(\boldsymbol{\phi} \,|\, \boldsymbol{y})$, as described earlier;
- Draw $\boldsymbol{\theta}^{(t)}$ from $p(\boldsymbol{\theta} \,|\, \boldsymbol{\phi}^{(t)}, \boldsymbol{y})$ as follows:
  - for $i \in \mathcal{R}(\boldsymbol{z}^{(t)})$, draw conditionally independent samples $\theta_i^{(t)}$ from

$$p(\theta_i^{(t)} \,|\, \boldsymbol{\phi}^{(t)}, y_i) = \mathcal{N}_{\mathrm{t}}\Big(\theta_i \,; \, \widehat{\theta}_i(\mu^{(t)}, \tau^{(t)}), \Big[\frac{1}{(\tau^{(t)})^2} + \frac{1}{\sigma^2}\Big]^{-1}\Big),$$

  - for $i \in \mathcal{R}^{\mathrm{c}}(\boldsymbol{z}^{(t)})$, set $\theta_i^{(t)} = 0$,

  yielding $\boldsymbol{\theta}^{(t)} = [\theta_1^{(t)}, \theta_2^{(t)}, \ldots, \theta_{\mathbb{N}_{\mathrm{tot}}}^{(t)}]^T$.

Computing the Mean Signal within $\mathcal{R}(z)$

Define the mean signal within the potential defect region:

$$\overline{\theta} = \frac{1}{N(\boldsymbol{z})} \cdot \sum_{i,\, \boldsymbol{s}_i \in \mathcal{R}(\boldsymbol{z})} \theta_i.$$

Then, $\overline{\theta}$ simulated in the tth draw is as

$$\overline{\theta}^{(t)} = [1/N(\boldsymbol{z}^{(t)})] \cdot \sum_{i,\, \boldsymbol{s}_i \in \mathcal{R}(\boldsymbol{z}^{(t)})} \theta_i^{(t)}.$$

Computing the Defect Area

Define the potential defect area to be proportional to the number of measurement locations i having signals θ_i that are within 10 dB from the maximum signal $\theta_{\mathrm{MAX}} = \max_{i,\, \boldsymbol{s}_i \in \mathcal{R}(\boldsymbol{z})} \theta_i$ in the potential defect region $\mathcal{R}(\boldsymbol{z})$.

Then, the area of the defect region simulated in the tth draw is

$$\text{defect area}^{(t)} \propto \text{ number of } \theta_i\text{s within 10 dB from } \theta_{\mathrm{MAX}}^{(t)}$$

where

$$\theta_{\mathrm{MAX}}^{(t)} = \max_{i,\, \boldsymbol{s}_i \in \mathcal{R}(\boldsymbol{z}^{(t)})} \theta_i^{(t)}.$$

Ranking Potential Defects Using Bayes Factors

Bayes factor for comparing models $\mathrm{H}_0 : \mu = 0$ (defect absent) versus the alternative $\mathrm{H}_1 : \mu > 0$ (defect present):

$$\mathrm{BF} = \frac{\overbrace{\prod_{i=1}^{\mathbb{N}_{\mathrm{tot}}} \mathcal{N}(y_i; 0, \sigma^2)}^{\text{noise-only distribution, according to } \mathrm{H}_0}}{\text{marginal distribution of the data in hand under } \mathrm{H}_1}$$

$\equiv$ *likelihood ratio test statistic* for testing H_0 versus H_1. Here,

- marginal distribution of the data in hand under H_1 $\equiv$ prior-weighted average of the likelihood (under H_1).

Ranking Potential Defects Using Bayes Factors

Bayes factor (up to a multiplicative constant):

$$\begin{aligned}
\mathrm{BF} &= \left[\int l(\boldsymbol{y} \,|\, \boldsymbol{\phi})\, \pi_{\boldsymbol{\phi}}(\boldsymbol{\phi})\, d\boldsymbol{\phi} \right]^{-1} \\
&= \int \frac{q(\boldsymbol{\phi})}{l(\boldsymbol{y} \,|\, \boldsymbol{\phi})\, \pi_{\boldsymbol{\phi}}(\boldsymbol{\phi})}\, p(\boldsymbol{\phi} \,|\, \boldsymbol{y})\, d\boldsymbol{\phi}
\end{aligned}$$

where $q(\boldsymbol{\phi})$ is an arbitrary pdf having support within the support of the posterior distribution $p(\boldsymbol{\phi} \,|\, \boldsymbol{y})$.

Bayes-Factor Computation

$$\mathrm{BF} \quad \approx \quad \frac{1}{T} \cdot \sum_{t=t_0+1}^{t_0+T} \frac{q(\boldsymbol{\phi}^{(t)})}{l(\boldsymbol{y} \mid \boldsymbol{\phi}^{(t)})\, \pi_{\boldsymbol{\phi}}(\boldsymbol{\phi}^{(t)})}.$$

A Choice of $q(\boldsymbol{\phi})$

$$q(\boldsymbol{\phi}) = q_{x_{0,1}}(x_{0,1}) \cdot q_{x_{0,2}}(x_{0,2}) \cdot q_d(d) \cdot q_A(A) \cdot q_\varphi(\varphi) \cdot q_\mu(\mu) \cdot q_\tau(\tau)$$

where

$$q_{x_{0,1}}(x_{0,1}) = \text{uniform}\Big(\widehat{x}_{0,1,\text{MIN}}(T), \widehat{x}_{0,1,\text{MAX}}(T)\Big)$$

$$q_{x_{0,2}}(x_{0,2}) = \text{uniform}\Big(\widehat{x}_{0,2,\text{MIN}}(T), \widehat{x}_{0,2,\text{MAX}}(T)\Big)$$

$$q_d(d) = \text{uniform}\Big(\widehat{d}_{\text{MIN}}(T), \widehat{d}_{\text{MAX}}(T)\Big),$$

$$q_A(A) = \text{uniform}\Big(\widehat{A}_{\text{MIN}}(T), \widehat{A}_{\text{MAX}}(T)\Big)$$

$$q_\varphi(\varphi) = \text{uniform}\Big(\widehat{\varphi}_{\text{MIN}}(T), \widehat{\varphi}_{\text{MAX}}(T)\Big)$$

$$q_\mu(\mu) = \text{uniform}\Big(\widehat{\mu}_{\text{MIN}}(T), \widehat{\mu}_{\text{MAX}}(T)\Big), \quad q_\tau(\tau) = \text{uniform}\Big(\widehat{\tau}_{\text{MIN}}(T), \widehat{\tau}_{\text{MAX}}(T)\Big).$$

A Choice of $q(\phi)$ (cont.)

Here

$$
\begin{aligned}
\widehat{x}_{0,1,\mathrm{MIN}}(T) &= \min\{x_{0,1}^{(t_0)}, x_{0,1}^{(t_0+1)}, \ldots, x_{0,1}^{(t_0+T)}\} \\
\widehat{x}_{0,1,\mathrm{MAX}}(T) &= \max\{x_{0,1}^{(t_0)}, x_{0,1}^{(t_0+1)}, \ldots, x_{0,1}^{(t_0+T)}\} \\
\widehat{x}_{0,2,\mathrm{MIN}}(T) &= \min\{x_{0,2}^{(t_0)}, x_{0,2}^{(t_0+1)}, \ldots, x_{0,2}^{(t_0+T)}\} \\
\widehat{x}_{0,2,\mathrm{MAX}}(T) &= \max\{x_{0,2}^{(t_0)}, x_{0,2}^{(t_0+1)}, \ldots, x_{0,2}^{(t_0+T)}\} \\
\widehat{d}_{\mathrm{MIN}}(T) &= \min\{d^{(t_0)}, d^{(t_0+1)}, \ldots, d^{(t_0+T)}\} \\
\widehat{d}_{\mathrm{MAX}}(T) &= \max\{d^{(t_0)}, d^{(t_0+1)}, \ldots, d^{(t_0+T)}\} \\
&\cdots
\end{aligned}
$$

Bayesian Analysis: Summary

- We developed *MCMC methods* for sampling from the posterior distributions of
  - the model parameters $\boldsymbol{\phi}$ and
  - random signals $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_{\mathbb{N}_{\mathrm{tot}}}]^T$.
- We utilize these samples to construct
  - minimum mean-square error (MMSE) estimates and
  - Bayesian confidence regions (credible sets)

  for ϕ and $\boldsymbol{\theta}$.

Estimation of ϕ and θ

Once we have collected enough samples, we estimate the posterior means of $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ simply by averaging the last T draws:

$$\mathrm{E}\left[\boldsymbol{\phi} \,|\, \boldsymbol{y}\right] \approx \widehat{\boldsymbol{\phi}} = \frac{1}{T} \sum_{t=t_0+1}^{t_0+T} \boldsymbol{\phi}^{(t)}, \quad \mathrm{E}\left[\boldsymbol{\theta} \,|\, \boldsymbol{y}\right] \approx \widehat{\boldsymbol{\theta}} = \frac{1}{T} \sum_{t=t_0+1}^{t_0+T} \boldsymbol{\theta}^{(t)}$$

where t_0 defines the burn-in period. Here, $\widehat{\boldsymbol{\phi}}$ and $\widehat{\boldsymbol{\theta}}$ are the (approximate) *MMSE* estimates of $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$.

Note: The proposed MCMC algorithms are *automatic*, i.e. their implementation *does not* require preliminary runs and additional tuning.

Outline

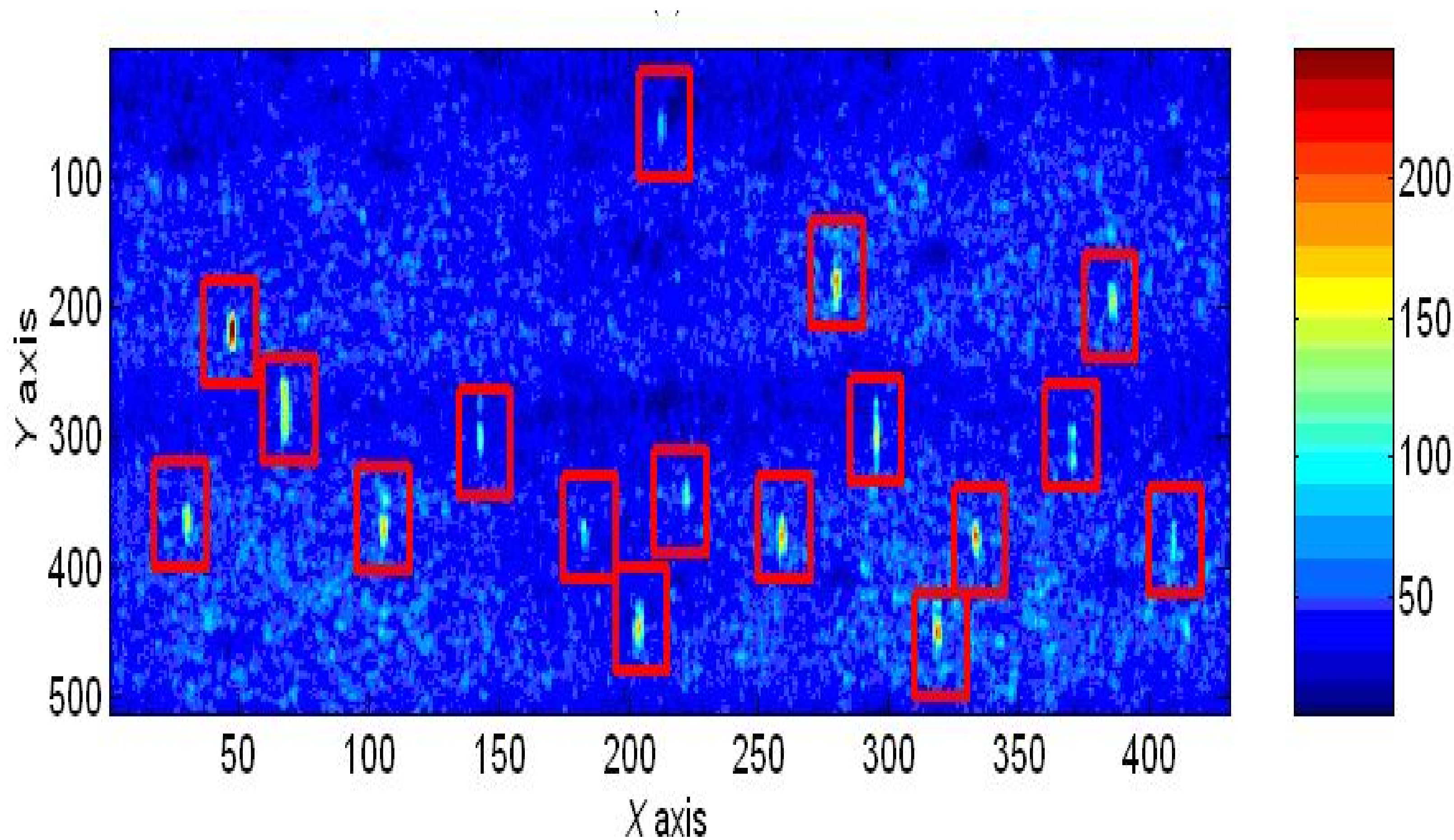
- Measurement Model
  - Parametric model for defect location and shape,
  - Measurement-error (noise) model,
  - Defect-signal (reflectivity) model,
  - Prior specifications for the location, shape, and defect-signal distribution parameters (model parameters).

- Bayesian Analysis
  - Simulating the model parameters $\boldsymbol{\phi}$,
  - Simulating the signals θ_i.

- Numerical Examples.

Numerical Example: Experimental UT Data

- Ultrasonic (UT) C-scan data from an inspection of a cylindrical Ti 6-4 billet.
- The sample, developed as a part of the work of the Engine Titanium Consortium, contains 17 $\#$ 2 flat bottom holes at 3.2" depth.
- The ultrasonic data were collected in a single experiment by moving a probe along the axial direction and scanning the billet along the circumferential direction at each axial position.
- The vertical coordinate is proportional to rotation angle and the horizontal coordinate to axial position.



Ultrasonic C-scan data with 17 defects.

Experimental UT Data Example

Prior Specifications:

$$
\begin{aligned}
\mu_{\text{MAX}} &= \max\{y_1, y_2, \ldots, y_{\mathbb{N}_{\text{tot}}}\} \\
\tau_{\text{MAX}} &= 7\,\sigma \\
d_{\text{MIN}} &= 1, \quad d_{\text{MAX}} = 10, \\
A_{\text{MIN}} &= 20, \quad A_{\text{MAX}} = 400 \\
\varphi_{\text{MIN}} &= -\pi/4, \quad \varphi_{\text{MAX}} = \pi/4, \qquad \text{(full range of } \varphi\text{s)}
\end{aligned}
$$

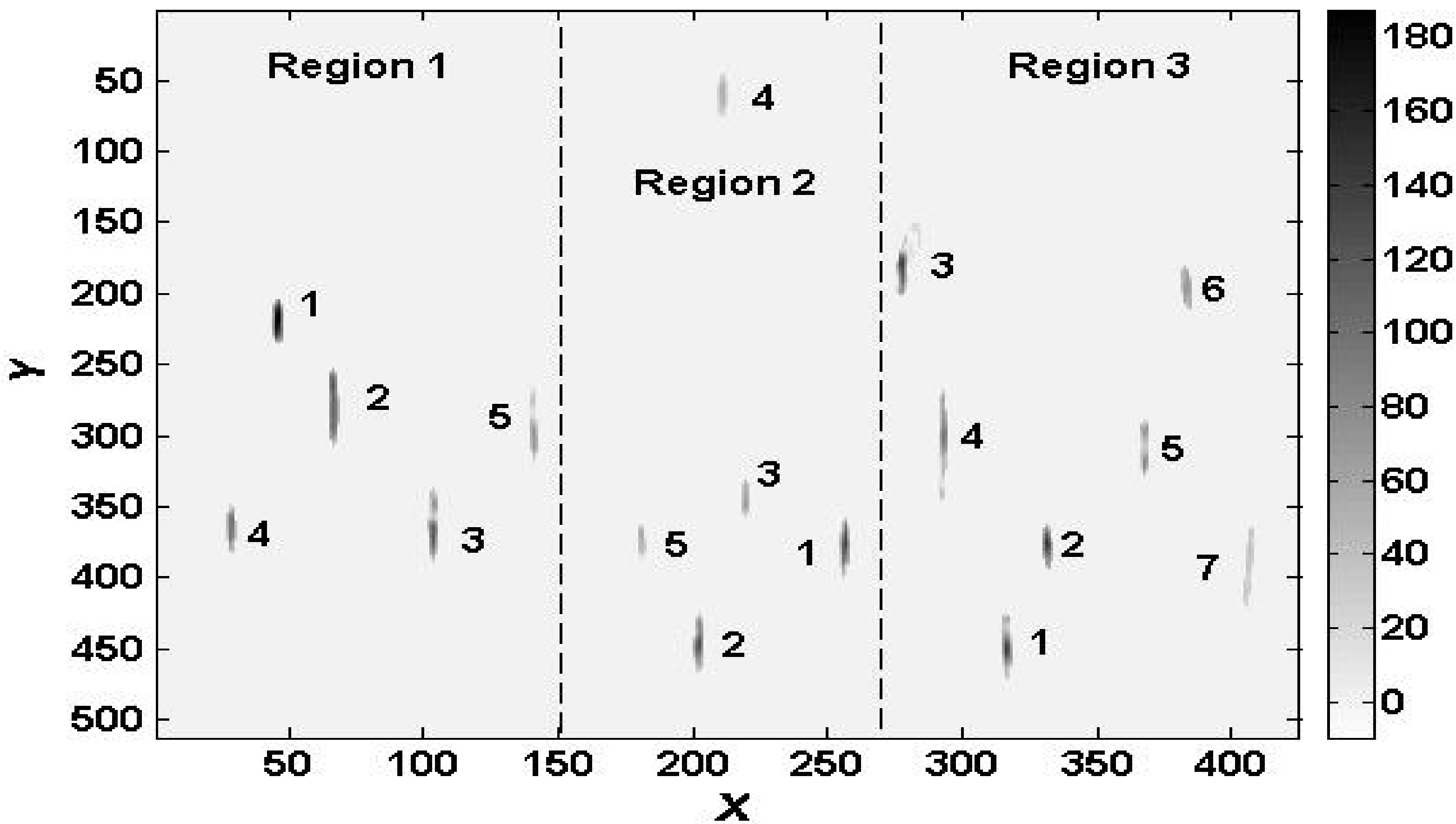
$x_{0,i,\text{MIN}}$,$x_{0,i,\text{MAX}}$, $i = 1, 2$ selected to span the entire region that is being analyzed.

Experimental UT Data Example (cont.)

- Before analyzing the data, we divided the C-scan image into 3 regions.
- In each region, we subtracted row means from the measurements within the same row.
- The noise level in Region 2 is lower than the corresponding noise levels in Regions 1 and 3. Indeed, the sample estimates of the noise variance σ^2 in Regions 1,2, and 3 are:

  $$\sigma^2 = 11.9^2, 10.3^2, \quad \text{and} \quad 12.0^2 \quad \text{(respectively)}.$$

  The underlying non-stationarity of the noise is due to the billet manufacturing process.
- In the following, we analyze each region separately assuming known noise variances σ^2, given above.



Regions 1,2, and 3.

Experimental UT Data Example (cont.)

We now describe our analysis of Region 1, where we ran seven Markov chains. We perform *sequential identification* of potential defects, as follows:

- Run $10,000$ cycles of the Gibbs sampler and utilized the last $T = 2,000$ samples to estimate the posterior pdfs $p(\boldsymbol{\phi} \,|\, \boldsymbol{y})$ and $p(\boldsymbol{\theta} \,|\, \boldsymbol{y})$.

- Get the approximate MMSE estimates of θ_i by averaging the T draws:

$$\widehat{\theta}_i\Big|_{\text{chain 1}} \approx \frac{1}{T} \sum_{t=t_0+1}^{t_0+T} \theta_i^{(t)}, \quad i = 1, 2, \ldots, N.$$

- *Subtract* the first chain's MMSE estimates $\widehat{\theta}_i\Big|_{\text{chain 1}}$ from the measurements y_i, $i = 1, 2, \ldots, N$, effectively removing the first potential defect region from the data. Then run the second Markov chain using the filtered data:

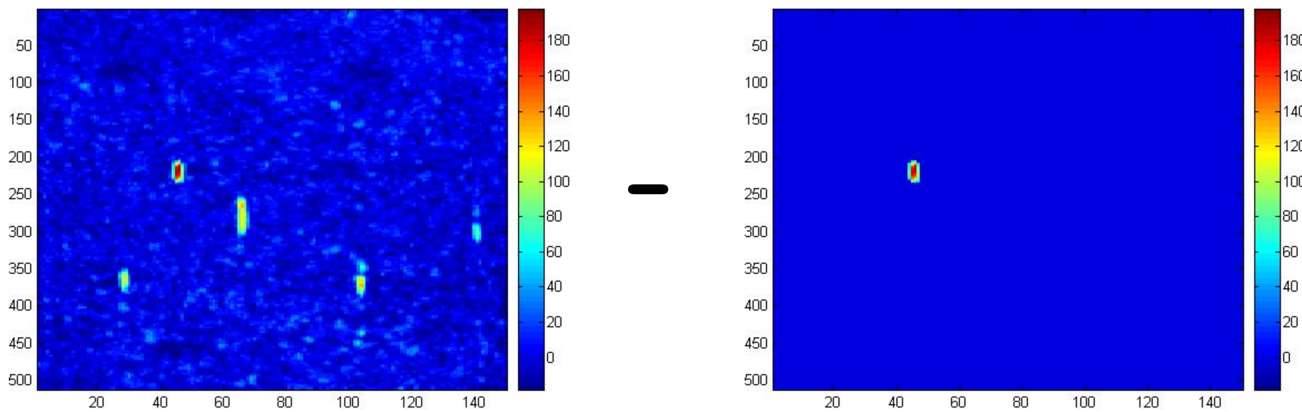
$$y_i\big|_{\text{chain 2}} = y_i - \widehat{\theta}_i\big|_{\text{chain 1}}$$

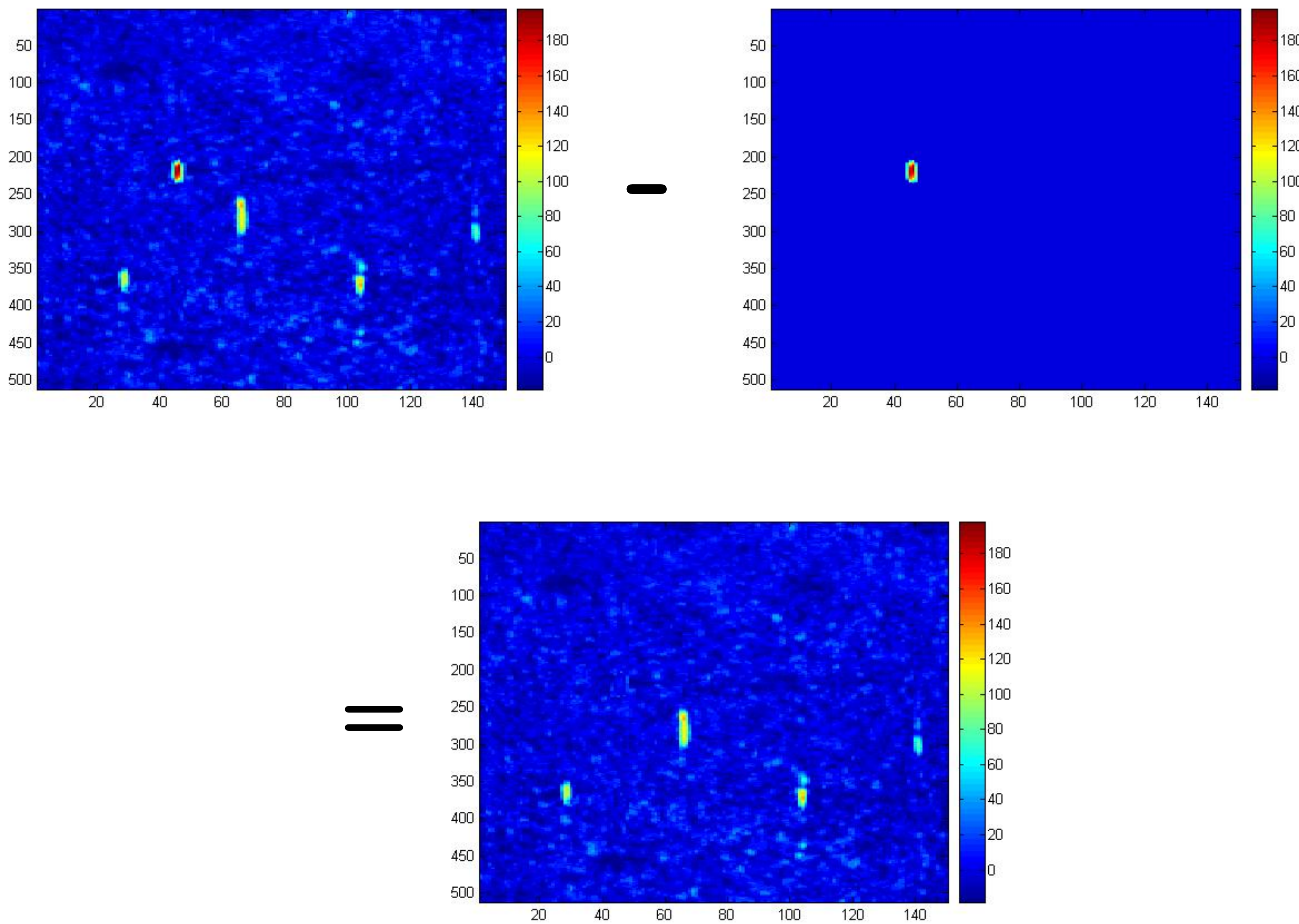
compute the approximate MMSE estimates $\widehat{\theta}_i\big|_{\text{chain 2}}$ of the second potential defect signal (using the second Markov chain), subtract them out, yielding

$$y_i\big|_{\text{chain 3}} = y_i\big|_{\text{chain 2}} - \widehat{\theta}_i\big|_{\text{chain 2}}$$

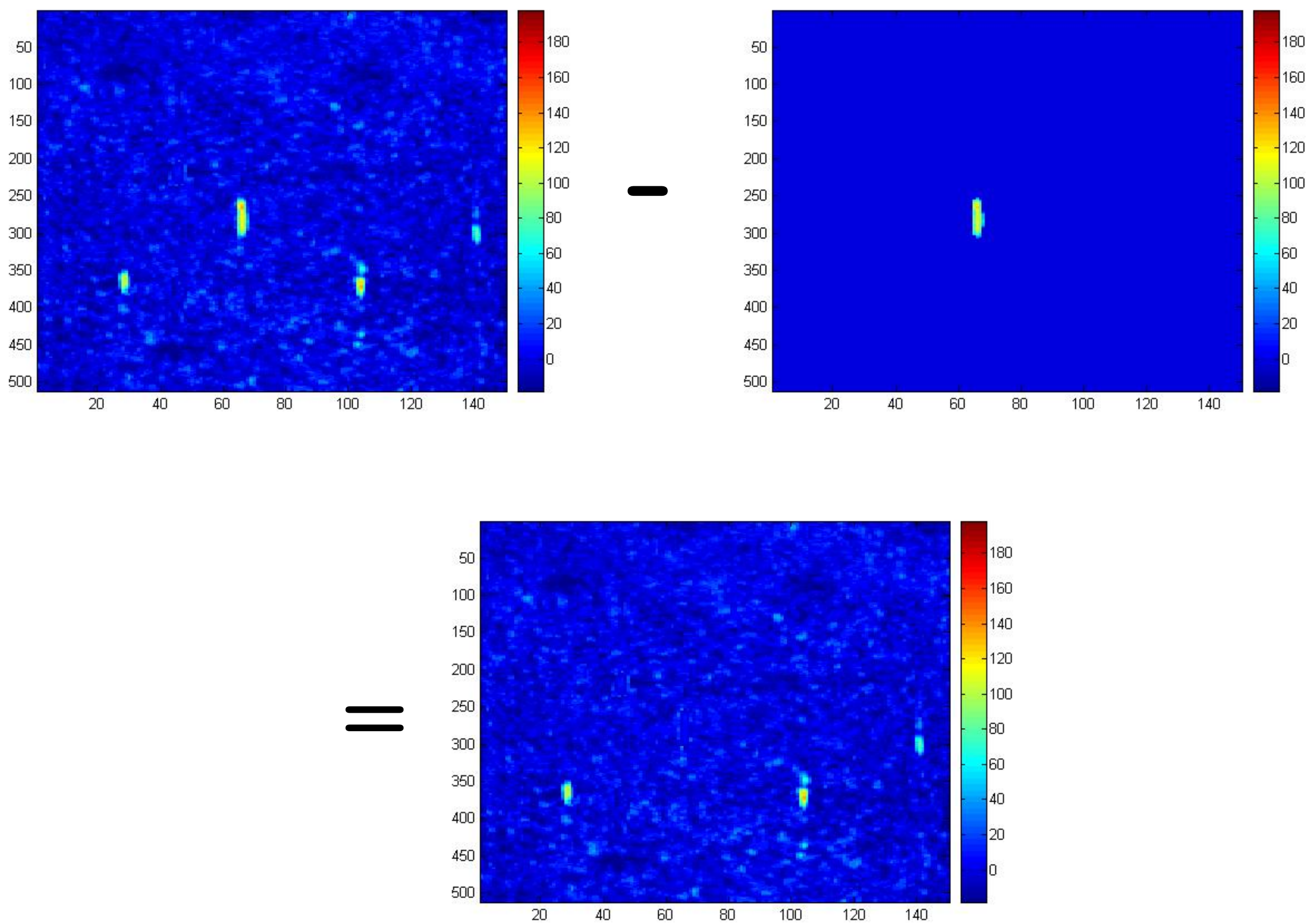
and continue this procedure until reaching the desired number of chains.

- We have applied the above sequential scheme to Regions 2 and 3 as well.

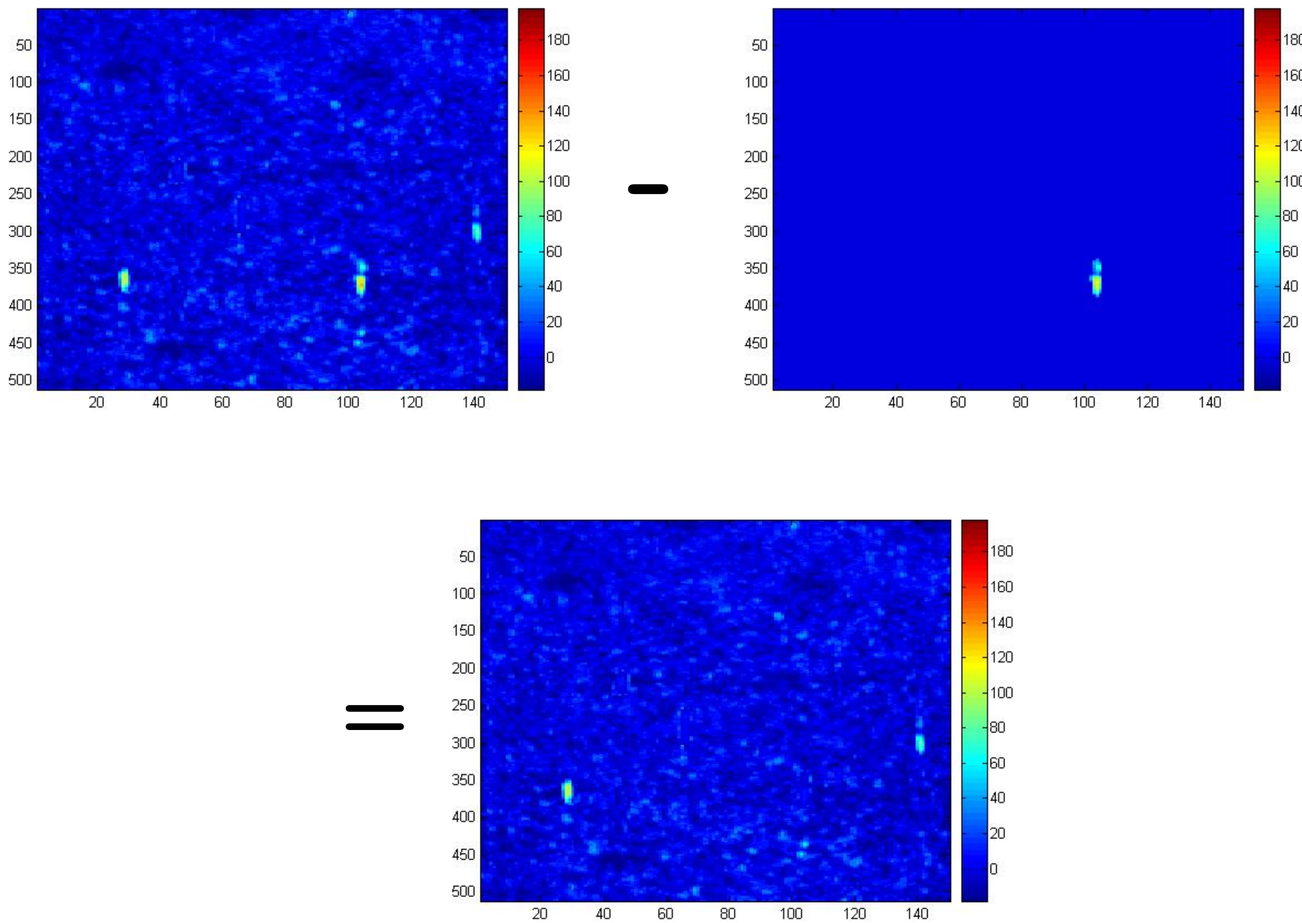

50
100
150
200
250
300
350
400
450
500
20
40
60
80
100
120
140
180
160
140
120
100
80
60
40
20
0
-
50
100
150
200
250
300
350
400
450
500
20
40
60
80
100
120
140
180
160
140
120
100
80
60
40
20
0



Our Sequential Identification in Action.



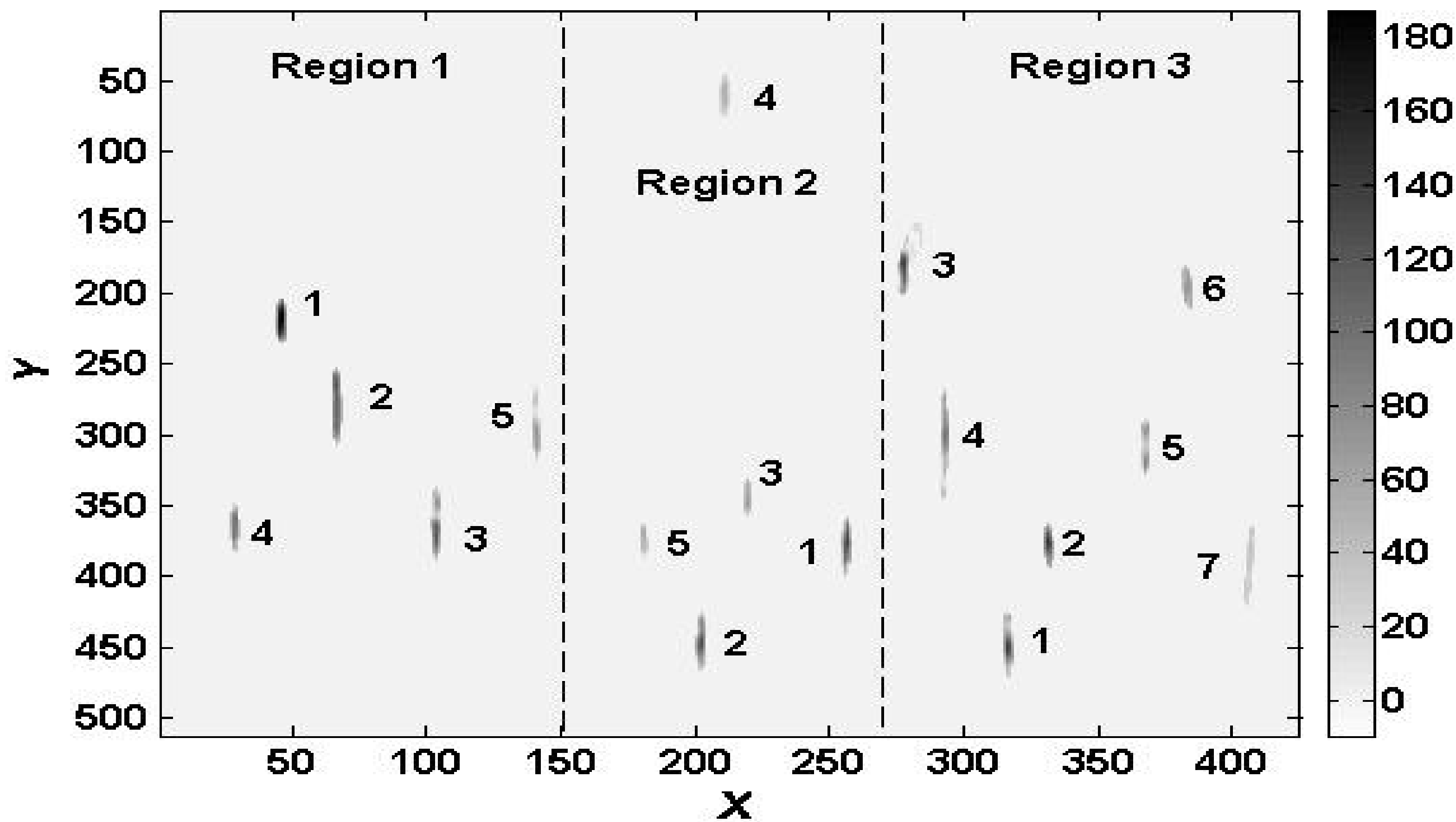
Our Sequential Identification in Action.



Our Sequential Identification in Action.

Our Sequential Identification in Action

Can you guess where the “erased” defects were located?



Defect estimation results: Approximate MMSE estimates of the defect signals for the chains having the largest average log posterior pdfs.

Bayesian Confidence Regions (Credible Sets)

90% Bayesian confidence regions for the normalized mean signals $\overline{\theta}/\sigma$ and defect areas

$$\left([\text{defect area}, \frac{\overline{\theta}}{\sigma}] - [\widehat{\text{defect area}}, \frac{\widehat{\overline{\theta}}}{\sigma}]\right) C^{-1} \left(\left[\begin{array}{c} \text{defect area} \\ \frac{\overline{\theta}}{\sigma} \end{array} \right] - \left[\begin{array}{c} \widehat{\text{defect area}} \\ \frac{\widehat{\overline{\theta}}}{\sigma} \end{array} \right] \right) \leq \xi \tag{2}$$

computed for all 24 potential defects in the three regions, where

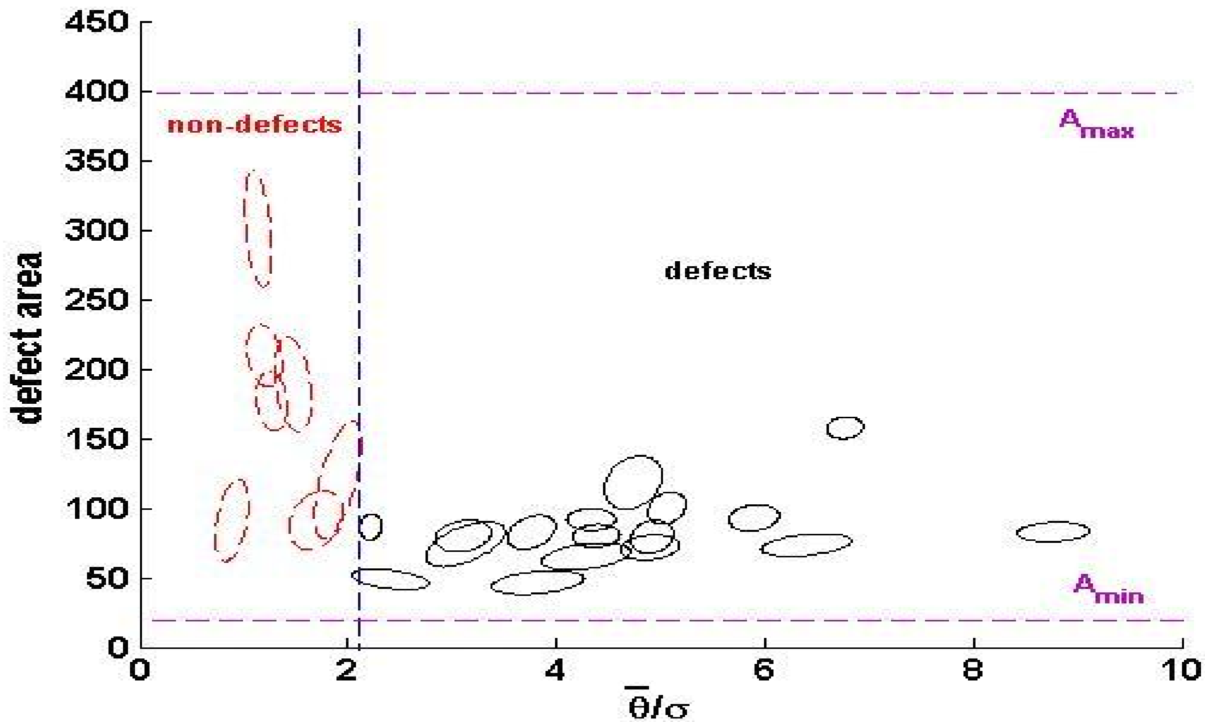
- $\widehat{\text{defect area}}$ and $\widehat{\overline{\theta}}$ denote the MMSE estimates of the defect area and $\overline{\theta}$,

- C is the sample covariance matrix of the posterior samples $[\text{defect area}^{(t)}, \overline{\theta}^{(t)}/\sigma]^T$:

$$
\begin{aligned}
C &= \frac{1}{T}\sum_{t=t_0+1}^{t_0+T}\left(\begin{bmatrix}\text{defect area}^{(t)}\\ \frac{\overline{\theta}^{(t)}}{\sigma}\end{bmatrix} - \begin{bmatrix}\widehat{\text{defect area}}\\ \frac{\widehat{\overline{\theta}}}{\sigma}\end{bmatrix}^T\right) \\
&\quad \cdot\left([\text{defect area}^{(t)}, \frac{\overline{\theta}^{(t)}}{\sigma}] - [\widehat{\text{defect area}}, \frac{\widehat{\overline{\theta}}}{\sigma}]\right)
\end{aligned}
$$

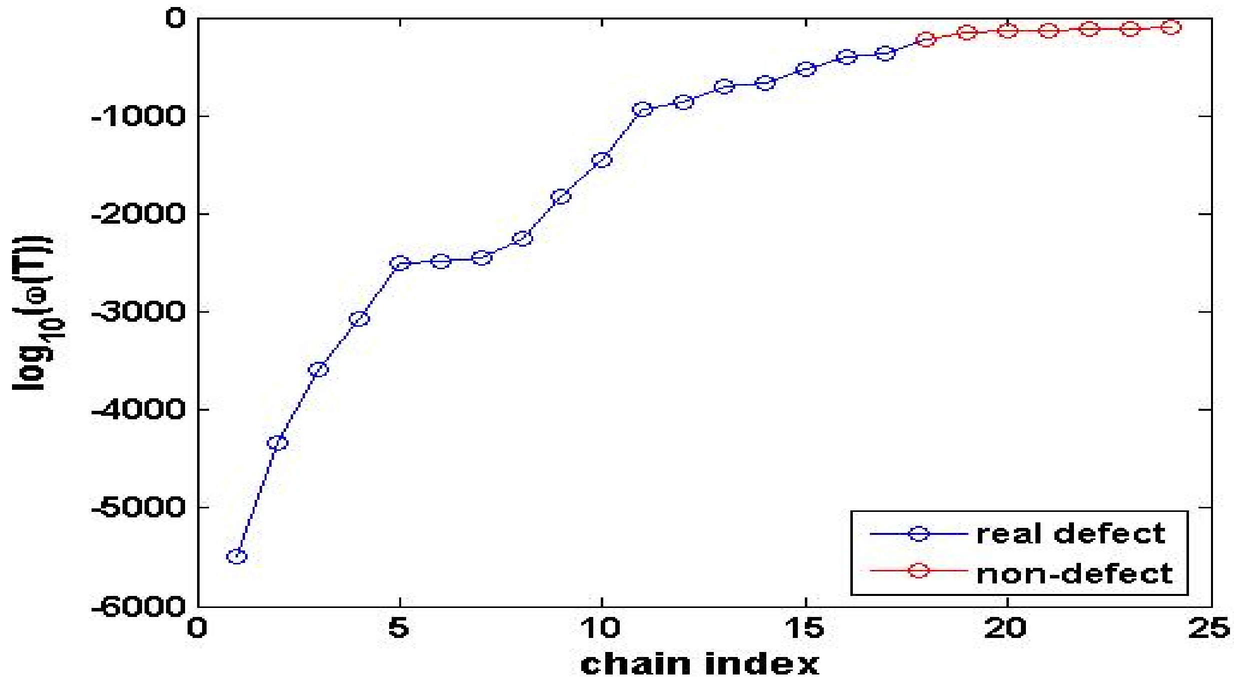
  and

- ξ is a constant chosen (for each chain) so that 90% of the samples $[\text{defect area}^{(t)}, \overline{\theta}^{(t)}/\sigma]^T$, $t = t_0, \ldots, T$ satisfy (2).



Approximate 90% credible sets for the normalized mean signals $\bar{\theta}/\sigma$ and areas A of all potential defects in the three regions and a possible classification boundary for separating defects from non-defects.

Ranking Potential Defects Using Bayes Factors

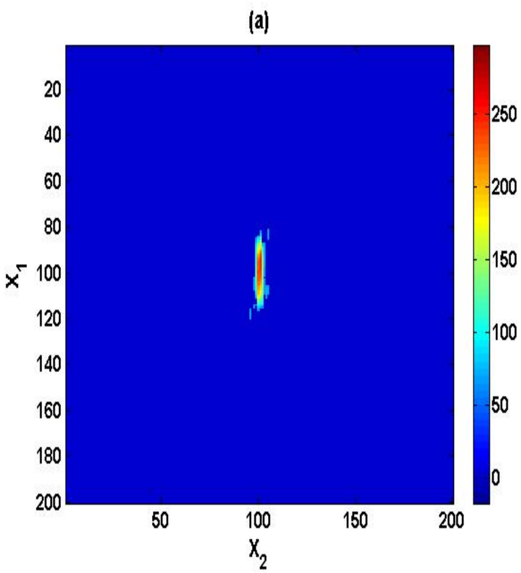


Logarithms of the estimated Bayes factors (up to an additive constant) for all 24 potential defects in the three regions of the inspected billet; non-defects are marked in red.

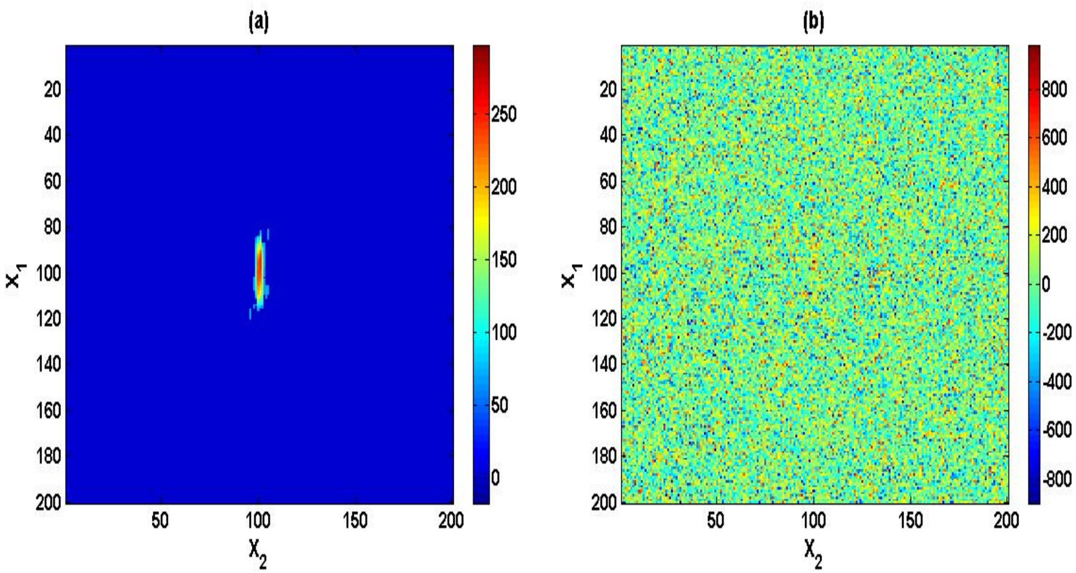
Simulated Example: Low-SNR Regime

- Low signal-to-noise ratio (SNR) scenario using simulated data. In particular,
  - we added i.i.d. zero-mean Gaussian noise with variance $\sigma^2 = 250^2$ to the defect signals θ_i coming from one of the flat bottom holes from the previous example.

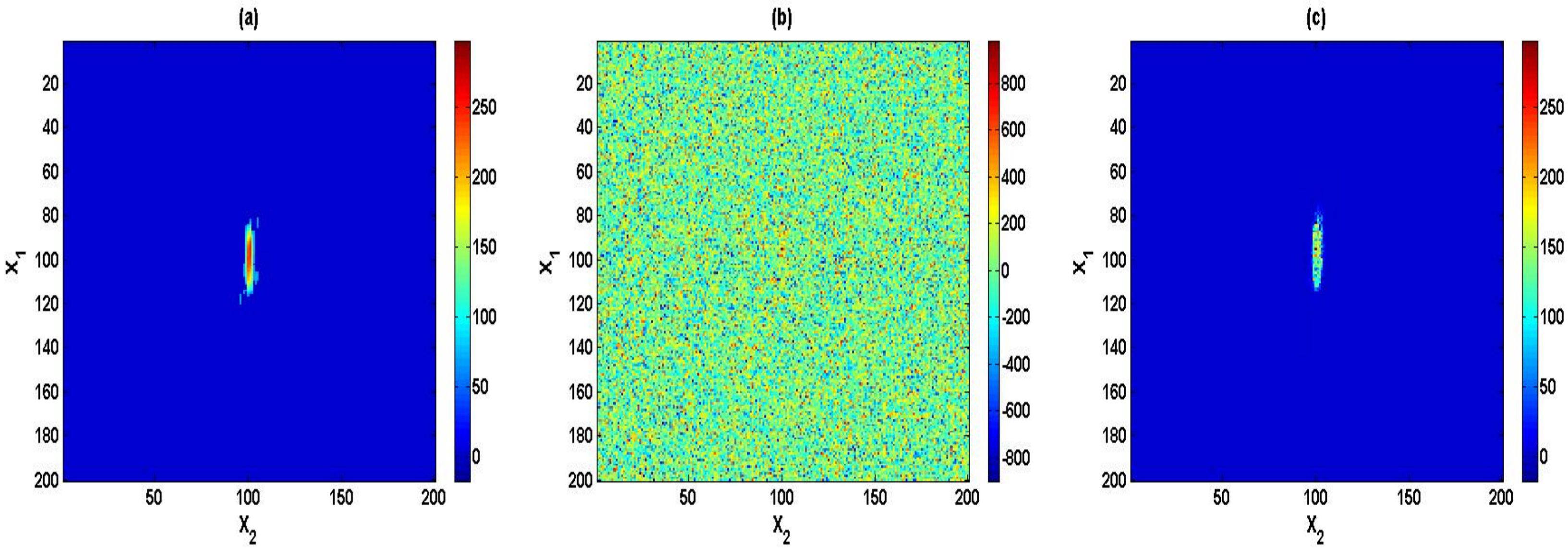
Simulated Example: Low-SNR Regime (cont.)



Simulated Example: Low-SNR Regime (cont.)



Simulated Example: Low-SNR Regime (cont.)



(a) Signals θ_i, (b) simulated noisy observations y_i, and (c) MMSE estimates $\widehat{\theta}_i$, for $i = 1, 2, \ldots, \mathbb{N}_{\text{tot}}$.

Summary of Accomplishments

Developed

- a parametric model that describes defect shape, location, and reflectivity,
- a hierarchical Bayesian framework and MCMC algorithms for estimating these parameters assuming a single defect,
- a sequential method for identifying multiple potential defect regions and estimating their parameters, and
- a simple classification scheme for separating defects from non-defects using estimated mean signals and areas of the potential defects.

Our approach provides Bayesian confidence regions (credible sets) for the estimated parameters $\Longrightarrow$ important in NDE applications.

Summary of Accomplishments (cont.)

The results of this work will be published in

A. Dogandžić and B. Zhang, "Markov chain Monte Carlo defect identification in NDE images," to appear in *Proc. Annu. Rev. Progress Quantitative Nondestructive Evaluation (QNDE 2006),* Portland, OR, Aug. 2006.

and, for the simpler case of Gaussian priors on θ_i within the defect region,

A. Dogandžić and B. Zhang, "Bayesian NDE defect signal analysis," *IEEE Trans. Signal Processing*, vol. 55, pp. 372–378, Jan. 2007.

Future Work

Developing Bayesian model-based methods that incorporate

- the forward model and
- realistic defect and noise models

into flaw detection, estimation (sizing), and system design.