Reversible Jump Markov Chain Monte Carlo for Pulse Fitting

Fred Goodyer Bashar Ahmad Simon Godsill Dept. of Engineering University of Cambridge

1. Problem

- Many industries must measure fluid levels: Time Domain Reflectometry (TDR).
- Electromagnetic pulse sent down probe; received reflections comprise echo curve. See **Figure 1**.
- Fluid level found from correct peak (here, τ_1).

Challenge: low fluid level ($\tau_1 \approx \tau_2$), as wrong peak at τ_2 (end of probe) blocks the correct one at τ_1.

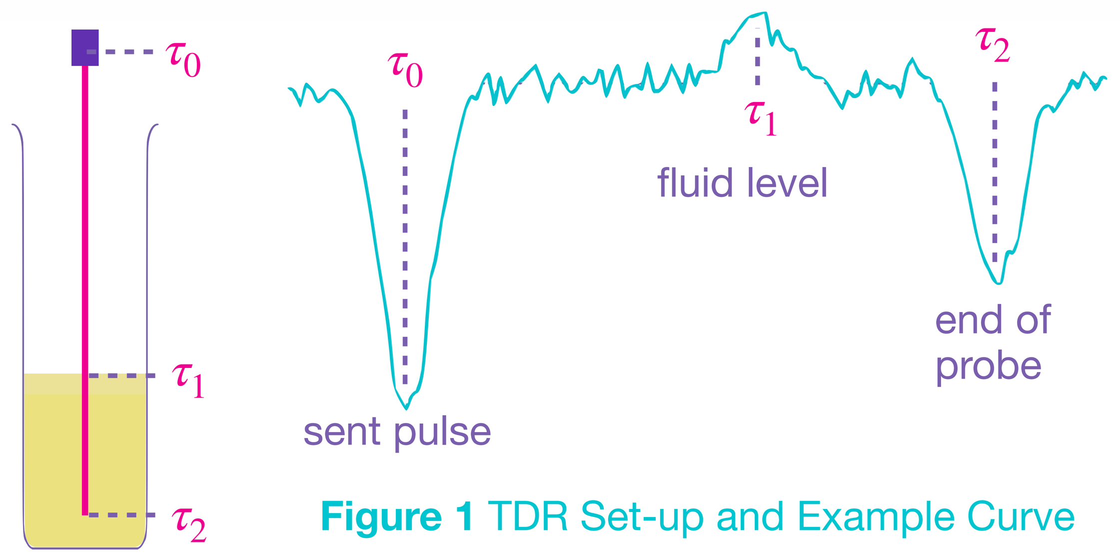

Figure 1 TDR Set-up and Example Curve

2. Modelling Assumptions

- Data y_k are observations of echo curve, $x(t)$, with Gaussian noise.

$$y_k = x(t_k) + \varepsilon_k, \quad \varepsilon_k \overset{iid}{\sim} \mathcal{N}(0, \sigma_y^2)$$

- Echo curve has ***unknown*** number of Gaussian peaks, N; peak i has ***unknown*** location, μ_i, and amplitude, A_i.

$$x(t) = \sum_{i=1}^{N} A_i \phi(t \mid \mu_i, \sigma^2), \quad \phi(t \mid \mu, \sigma^2) = e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$

- Form of ϕ, σ^2 and σ_y^2 being ***known*** is easily generalised.

3. RJ-MCMC

Method: Reversible Jump Markov Chain Monte Carlo (RJ-MCMC).

- Proposal $j \le S$ adds or removes a peak, or adjusts the current peaks.
- Sample step type by $\tilde{N}^j \sim q_N(N \mid N^{j-1})$.

Adjust ($\tilde{N}^j = N^{j-1}$)

Peaks are adjusted sequentially. For peak i:

- Location sampled $\tilde{\mu}_i^j \sim q_\mu(\mu_i \mid \boldsymbol{\mu}_{-i}^j, \mathbf{A}_{-i}^j, \mathbf{y}, \tilde{N}^j)$
- Amplitude sampled $\tilde{A}_i^j \sim p(A_i \mid \mathbf{A}_{-i}^j, \tilde{\boldsymbol{\mu}}_i^j, \mathbf{y}, \tilde{N}^j)$

where $\mathbf{A}_{-i} = \mathbf{A} \backslash A_i$, and similarly for $\boldsymbol{\mu}_{-i}$.
Each $(\tilde{\mu}_i, \tilde{A}_i)$ accepted with tractable probability α_i^j.

Add Peak ($\tilde{N}^j = N^{j-1} + 1$)

- Location sampled $\tilde{\mu}_b^j \sim q_\mu(\mu_b \mid \boldsymbol{\mu}^{j-1}, \mathbf{A}^{j-1}, \mathbf{y}, \tilde{N}^j)$
- ***All*** amplitudes sampled $\tilde{\mathbf{A}}^j \sim p(\mathbf{A} \mid \tilde{\boldsymbol{\mu}}_b^j, \mathbf{y}, \tilde{N}^j)$.

Sample $(\tilde{\mu}_b^j, \tilde{\mathbf{A}}^j)$ accepted with probability β^j.

Remove Peak ($\tilde{N}^j = N^{j-1} - 1$)

- Deleted peak sampled $I^j \sim p(i \mid \mathbf{A}^{j-1}, \mathbf{y}, \tilde{N}^j)$
- ***All*** amplitudes sampled $\tilde{\mathbf{A}}^j \sim p(\mathbf{A} \mid \tilde{\boldsymbol{\mu}}_{-I^j}^{j-1}, \mathbf{y}, \tilde{N}^j)$.

Both I^j and $\tilde{\mathbf{A}}^j$ jointly accepted with probability γ^j.

Scheme Summary: see **Algorithm**.

Paper

Code

Task

Learn:

- number of peaks, N
- all peaks, $\{\mu_i, A_i\}_{i=1}^N = \{\boldsymbol{\mu}, \mathbf{A}\}$

Contribution

Peak Proximity Parameter (**PPP**), δ, can drastically improve efficiency and accuracy.

4. Priors

- Number of peaks: $N \sim \text{Po}(\lambda)$
- Vector of amplitudes: $\mathbf{A} \mid N \sim \mathcal{N}(\mathbf{m}_A, \Sigma_A)$
- Vector of locations: $\boldsymbol{\mu} \mid N \sim U(S_N)$:

with $S_N = \{\boldsymbol{\mu}' : |\mu_i' - \mu_{i'}'| > \delta \ \forall \, i, i' \le N\}$.

In words: locations are uniform, but no two peaks are closer than δ, the **PPP**.

Algorithm

for $1 \le j \le S$:
  sample step type $\tilde{N}^j \sim q_N(N \mid N^{j-1})$
  if $\tilde{N}^j = N^{j-1}$:
    for $1 \le i \le \tilde{N}^j$:
      sample $\tilde{\mu}_i^j, \tilde{A}_i^j$; $\mathbb{P}(\text{accept}) = \alpha_i^j$
  elif $\tilde{N}^j = N^{j-1} + 1$:
    sample $\tilde{\mu}_b^j, \tilde{\mathbf{A}}^j$; $\mathbb{P}(\text{accept}) = \beta^j$
  elif $\tilde{N}^j = N^{j-1} - 1$:
    sample $I^j, \tilde{\mathbf{A}}^j$; $\mathbb{P}(\text{accept}) = \gamma^j$

Output: $\{\boldsymbol{\mu}^j, \mathbf{A}^j, N^j\}_{j=B+1}^S$

5. Proposal Distributions

Number of Peaks

Choose $q_N(1 \mid 0) = 1$, and for $N > 0$, choose $q_N(N \pm 1 \mid N) = c$, and $q_N(N \mid N) = 1 - 2c$.

Locations

For time steps $t_k : |t_k - \mu_i| > \delta, \forall \, i$, use discrete convolution of ϕ and residuals r_k:

$$q_\mu(t_k \mid \boldsymbol{\mu}, \mathbf{A}, \mathbf{y}, N) \propto \left| \sum r_{k'} \phi(t_{k'} \mid t_k, \sigma^2) \right|$$

For $\hat{t}_k : |\hat{t}_k - \mu_i| = \delta$ for some μ_i, set q_μ to zero:

$$q_\mu(\hat{t}_k \mid \boldsymbol{\mu}, \mathbf{A}, \mathbf{y}, N) = 0$$

Rest of q_μ uses linear interpolation: see **Figure 2**.

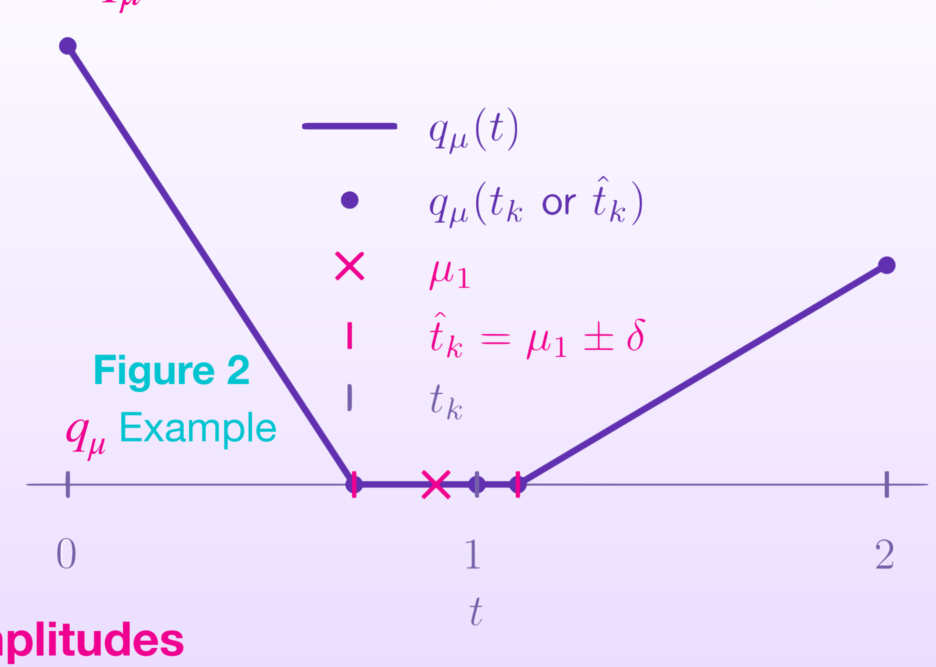

Figure 2 q_μ Example

Amplitudes

Conditional posterior:

$$p(\mathbf{A} \mid \boldsymbol{\mu}, \mathbf{y}, N) = \mathcal{N}(\mathbf{A} \mid \tilde{\mathbf{m}}_A, \tilde{\Sigma}_A).$$

Removal Index

Choose $p(i \mid \mathbf{A}, N) \propto |A_i|^{-1}$.

6. Experiments

Synthetic Data

- Run on 100 data sets with and without the **PPP**.
- Ideally, using δ learns curve better and faster.
- **Table 1** $\Rightarrow$ Using δ better on all counts.
- **Figures 3** & **4** show the better fitted curves and estimation of N for data set #100.

Synthetic Data Metrics

- Root Mean Square Error (RMSE): error between learnt and true curves
- Overlap: area under each pair of same-sign peaks
- Cancel: $1 - \frac{\text{area under |sample curve|}}{\sum_i |A_i^j|}$
- N: number of peaks
- Time (s): run time

'Overlap' and 'Cancel' measure overfitting.

	δ	No δ	Truth
RMSE	**3.88**	4.31	0
Overlap	**0.04**	0.15	0.04
Cancel	**0.07**	0.39	0.11
N	**4.39**	12.4	4.86
Time	**2.25**	6.40	

Table 1 Avg. Synthetic Results

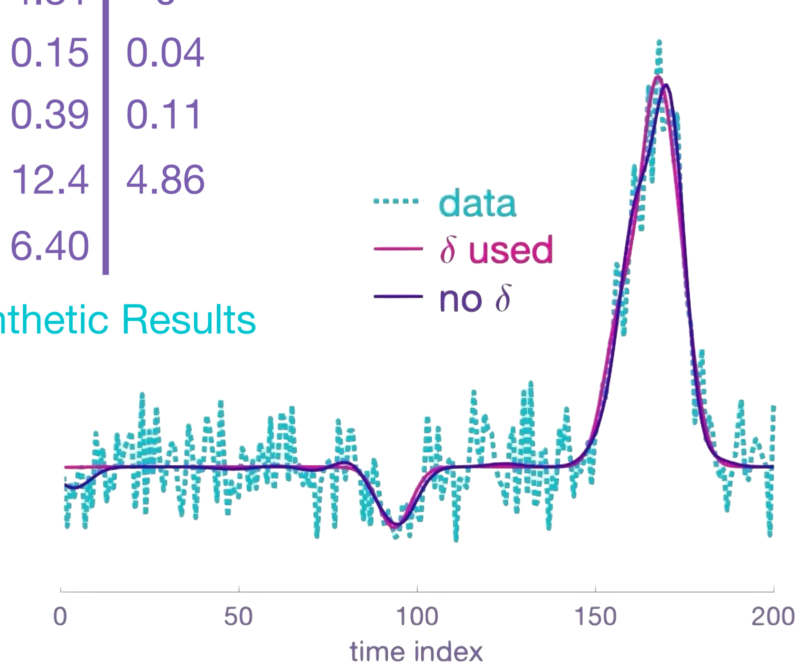

Figure 3 Data Set #100 – Fitted Curves

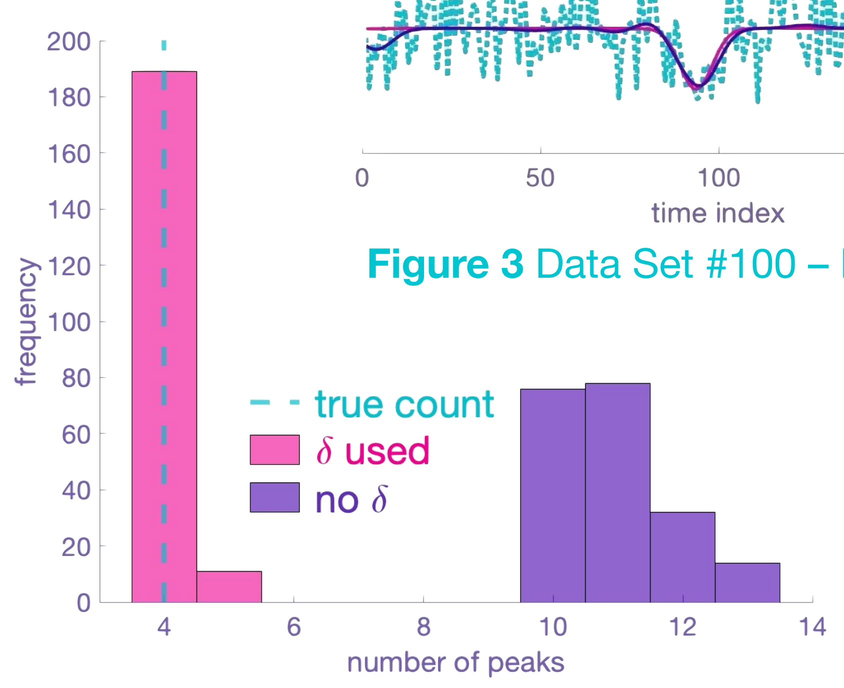

Figure 4 Data Set #100 – Posterior N Samples

Real Data (from SICK AG.)

- Low level of oil makes peak at 113.5.

Challenge: heavy overlap with end of probe peak.

- **Figures 5** & **6** show every peak from every sample for each method.
- **Table 2** $\Rightarrow$ Good estimate without δ; better and faster with δ.

	δ	No δ	Truth
Oil	**111.3**	109.9	113.5
Error	**2.2**	3.6	0
N	**2.4**	4.5	
Time	**1.0**	2.8	

Table 2 Real Data Results

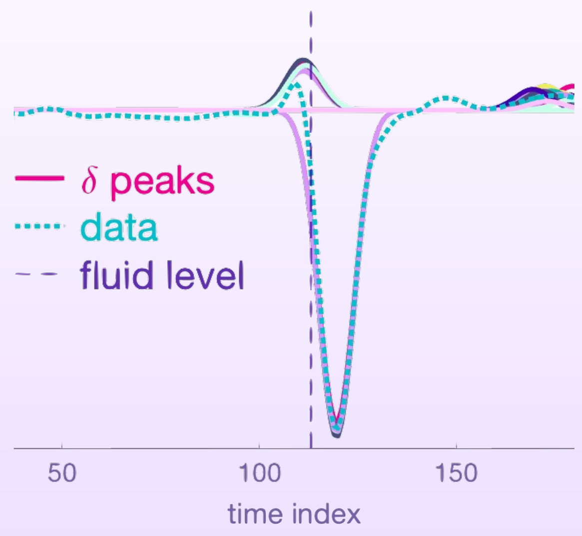

Figure 5 Real Data – All Peaks with δ

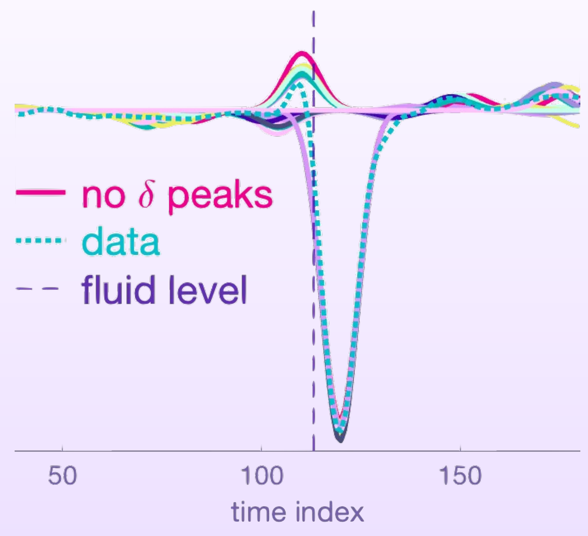

Figure 6 Real Data – All Peaks without δ

UNIVERSITY OF CAMBRIDGE