Parameter Estimation in Polymerization Systems Using Markov Chain Monte Carlo Methods

M. Mathew, T. Duever, IPR Symposium, University of Waterloo, ON N2L 3G1, Canada

Classical parameter estimation techniques such as non-linear least squares and maximum likelihood have been extensively applied to non-linear models. However, one limitation in using such methods is that they often involve optimization techniques that may only converge to a local minimum instead of the desired global minimum. In addition, capturing parameter uncertainty requires applying formulas derived from linear regression, resulting in approximate confidence regions. Both these issue can be addressed using Markov Chain Monte Carlo methods; a technique that not only provides reliable parameter point estimates but also produces joint confidence regions (JCRs) with exact shape and exact probability content. In light of these advantages, MCMC methods will be applied to estimate reactivity ratios in the copolymerization of di-n-butyl itaconate (DBI) and methyl methacrylate (MMA). The reactivity ratios will be determined using both the Mayo-Lewis and Meyer Lowry Models, and the data will be analyzed using MCMC methods and compared to classical regression approaches.

Markov Chain Monte Carlo Technique

The goal of parameter estimation is to find parameter values that minimize the difference between the predicted and experimentally observed values. To accomplish this, Markov Chain Monte Carlo starts by randomly drawing samples from the posterior probability distribution function. Since drawing samples with the correct frequency from this probability function can be problematic, there are a variety of MCMC sampling methods available. The method used in this study is known as the Metropolis Hastings algorithm, which was first proposed by Metropolis and Rosenbluth (1953) and later generalized by Hastings (1970).

After a sufficient amount of samples are generated, parameter point estimates can be determined by taking the expected value of the samples; a technique that is referred to as Monte Carlo integration. Therefore, by using MCMC methods, complex optimization algorithms can be replaced by a simple Monte Carlo integration.

In addition to parameter point estimates, the samples can also be used to construct 95% joint confidence regions. Since the samples are being drawn from the exact probability density function, the JCRs will have exact shape and exact probability content.

Case Study: Di-n-Butyl Itaconate/Methyl Methacrylate

The free radical copolymerization of di-n-butyl itaconate (DBI) and methyl methacrylate (MMA), conducted by Madruga and Fernandez-Garcia (1994), will be used as a case study to illustrate the implementation and advantages of MCMC methods. The free radical copolymerization was carried out in benzene at 50ºC and at low conversions levels. The data obtained from the experiment will be studied using both the Mayo-Lewis and Meyer Lowry models.

Reactivity Ratio from Mayo-Lewis Model

The Mayo-Lewis model, shown in equation 1, provides a relationship between the instantaneous copolymer composition, the monomer feed composition and the reactivity ratios. The model is only valid, however, at low conversion levels due to the inherent assumption that the compositional drift in the monomer feed is negligible.

$$F_1 = \frac{r_1 f_1^2 + f_1 f_2}{r_1 f_1^2 + 2 f_1 f_2 + r_2 f_2^2} \tag{1}$$

Using the copolymerization data and the above model, the Metropolis-Hastings algorithm was executed for 1,000,000 cycles. The average of the samples was taken to determine the reactivity ratio point estimate and 95% joint confidence region was constructed. The JCRs from the MCMC algorithm were then compared to the JCRs in Kazemi et al., (2011) where a linear regression approach was employed. The findings are summarized in figure 1.

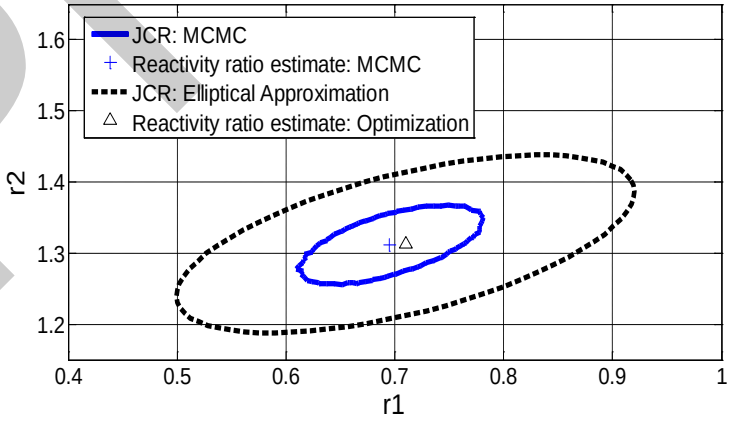


Figure 1: Point estimates and joint confidence regions for the reactivity ratios using the Mayo-Lewis Model in copolymerization of DBI/MMA. The data was analyzed using both MCMC and classical regression approaches.

Although point estimates are important, they need to be complimented by the corresponding JCRs in order to fully understand the system. The joint confidence region arises from the uncertainty in the data and it represents all the possible values of r_1 and r_2. Comparison of the two approaches in figure 1 shows that although both methods provides JCRs with similar shape, the linear regression approach overestimate the uncertainty in the reactivity ratios. When using the approximate JCRs, a significant amount of r_1 and r_2 values are included when they should be excluded based on a confidence level of 95%.

Different JCRs for both methods also translate to a significant difference in the Mayo-Lewis plot. In figure 2, the middle curve represents the Mayo-Lewis plot obtained using the reactivity ratio point estimates. The other curves to the right and left of this line indicate the upper and lower limits for each method, obtained by taking lowest and highest points in both JCRs.

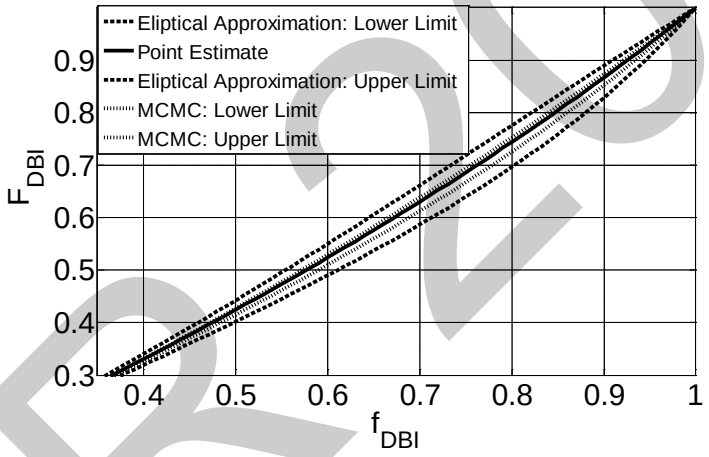


Figure 2: The Mayo-Lewis plot for the copolymerization of DBI/MMA, using the final molar composition, F_{DBI}, and the initial monomer feed composition, f_{DBI}. The five curves represent how the Mayo-Lewis relationship changes at different uncertainty levels.

Figure 2 shows significant differences between the two approaches, especially at a higher mole fraction of DBI. This example further demonstrates the importance of accurately capturing uncertainty in non-linear models and provides strong motivation for applying MCMC methods to non-linear models

Reactivity Ratio from Meyer-Lowry Model

The Mayo-Lewis equation was integrated by Meyer and Lowry (1965) and the analytical solution is known as the Meyer-Lowry Model, shown in equation 2. The assumption of negligible

compositional drift is no longer an issue since the Meyer-Lowry Model taken into account the conversion of the polymerization reaction. This model can therefore be applied to data obtained at low to moderate conversion levels.

$$X_N = 1 - \left(\frac{f_{10} - \bar{F}_1 X_n}{f_{10}(1 - X_n)}\right)^{\alpha} \left(\frac{1 - X_n - f_{10} - \bar{F}_1 X_n}{(1 - f_{10})(1 - X_n)}\right)^{\beta} * \left(\frac{(\delta - f_{10})(1 - X_n)}{\delta - \delta X_n - f_{10} + \bar{F}_1 X_n}\right)^{\gamma} \quad (2)$$

Where, $\alpha = \frac{r_2}{(1-r_2)}$, $\beta = \frac{r_1}{(1-r_1)}$, $\gamma = \frac{1-r_1 r_2}{(1-r_1)(1-r_2)}$ $\delta = \frac{(1-r_1)}{(2-r_1-r_2)}$

In equation 2, the variables $\bar{F}_1$, X_n and f_{10} represents the cumulative copolymer composition, molar conversion and the initial monomer mole fraction respectively. Implementation of the MCMC technique for the Meyer-Lowry Model is slightly more challenging since equation 2 is in implicit form. Therefore, a MCMC algorithm must be coupled with a root finding algorithm that solves for $\bar{F}_1$. Due to computational limitation, Metropolis-Hastings algorithm was only executed for 100,000 cycles. The reactivity ratio estimates and confidence regions are once again summarized in Figure 3.

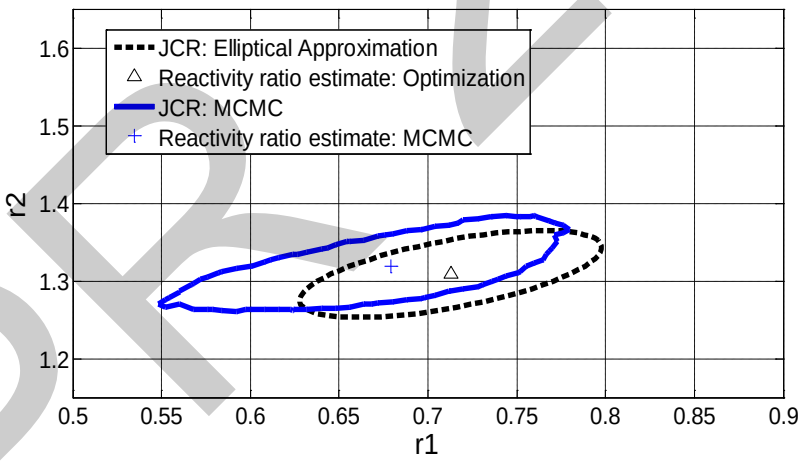


Figure 3: Point estimates and joint confidence regions for the reactivity ratios using the Meyer-Lowry Model in copolymerization of DBI/MMA. The data was analyzed using both MCMC and classical regression approaches.

The MCMC results shown in figure 3 indicate a confidence region that is not elliptical due to the non-linear nature of the model. It is also evident from figure 3 that using approximate JCRs will not capture the uncertainty in the parameter estimates accurately. The results are yet another example of the effectiveness of MCMC, especially when the model is highly non-linear.

References

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika, 57*(1), 97-109.

Kazemi, N., Duever, T.A, Penlidis, A., (2011). Reactivity Ratio Esimtation from Cumulative Copolymer Composition Data. *Macromolecular Reaction Engineering 5(9),* 385-403

Madruga, E., Fernandez-Farcia, M., (1994). Free-radical homopolymerization and copolymerization of di-n-butyl Itaconate. *Polymer, 35(20),* 4437–4442

Mayo, E.R., Lewis, F.M., (1944). Copolymerization. A basis for comparing the behavior of monomers in copolymerization; the copolymerization of Styrene and Methyl Methacrylate. *Journal of the American Chemical Society, 66(9),* 1594-1601

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics, 21*, 1087.

Meyer, V.E., Lowry, G.G., (1965). Integral and differential binary copolymerization equations. *Journal of Polymer Science, 3(8),* 2843–2851

Parameter Estimation in Polymerization Models using Markov Chain Monte Carlo Methods

Manoj Mathew

Supervisor: Professor Tom Duever

IPR Presentation

May 8, 2013

Outline

- Introduction
  - Non-Linear Regression
  - Classical Approaches and their Limitations
- Markov Chain Monte Carlo (MCMC)
  - Metropolis Hastings
- Polymerization Case Study
  - Mayo-Lewis Model Results
  - Meyer-Lowry Model Results
- Future Work
- Conclusions

Non-Linear Regression Approach

Experimental Data

Experimental Data

•Data ideally from designed experiments

•Can be multi-input and multi-response data

Non-Linear Regression Approach

Experimental Data → Model

Model

•Normally, error in only the dependent variable (y) is considered

•Error in Variables Model (EVM) takes into account error in all the variables

$y_i = y_i^* + \varepsilon_{yi}$

$x_i = x_i^* + \varepsilon_{xi}$

$y_i^* = f(x_i,\theta)$

$y_i^* = f(x^*,\theta)$

•Can be extended to multi-response multi-input models

Non-Linear Regression Approach

Experimental Data → Model → Measurement Error

Measurement Error

•Assume the error is normally distributed

•Assume constant error for all the experimental runs

$$\begin{pmatrix} \varepsilon_{xi} \\ \varepsilon_{yi} \end{pmatrix} \sim N\,(0, V\,)$$

$$V = \begin{pmatrix} \sigma_x{}^2 & 0 \\ 0 & \sigma_y{}^2 \end{pmatrix}$$

Non-Linear Regression Approach

Experimental Data → Model → Measurement Error → Bayesian Approach

Bayesian Approach

•A Bayesian approach allows us to use prior parameter knowledge

$$P(\theta|D) \propto P(\theta) * P(D|\theta)$$

Posterior Distribution — Prior Distribution — Likelihood Function

Non-Linear Regression Approach

Experimental Data → Model → Measurement Error → Bayesian Approach

$$P(\theta|D) \alpha P(\theta) * P(D|\theta)$$

Posterior Distribution | Prior Distribution | Likelihood Function

Prior Distribution

•A non-informative uniform prior is used when no prior information is known

Likelihood Function

$$L(\theta, x_i^*|y_i, x_i) \alpha \exp\left\{\frac{-1}{2}\sum_{i=1}^{n}\begin{pmatrix} x_i - x_i^* \\ y_i - f(x_i^*, \theta)\end{pmatrix} V^{-1}\begin{pmatrix} x_i - x_i^* \\ y_i - f(x_i^*, \theta)\end{pmatrix}\right\}$$

Non-Linear Regression Approach

Experimental Data → Model → Measurement Error → Bayesian Approach → Optimization

Optimization

•Parameter estimation involves finding the parameter values that maximize the posterior distribution function

$$D(\theta, x_i^*|y_i, x_i) \alpha \exp\left\{\frac{-1}{2}\sum_{i=1}^{n}\begin{pmatrix} x_i - x_i^* \\ y_i - f(x_i^*, \theta)\end{pmatrix} V^{-1}\begin{pmatrix} x_i - x_i^* \\ y_i - f(x_i^*, \theta)\end{pmatrix}\right\}$$

Limitation

The optimization may find a local minimum instead of the desired global minimum

Joint Confidence Regions

Elliptical Approach

$$(\theta - \hat{\theta})^T X^T X(\theta - \hat{\theta}) \leq ps^2 F_{1-\alpha}(p, v)$$

Wrong Shape And Wrong Probability Content

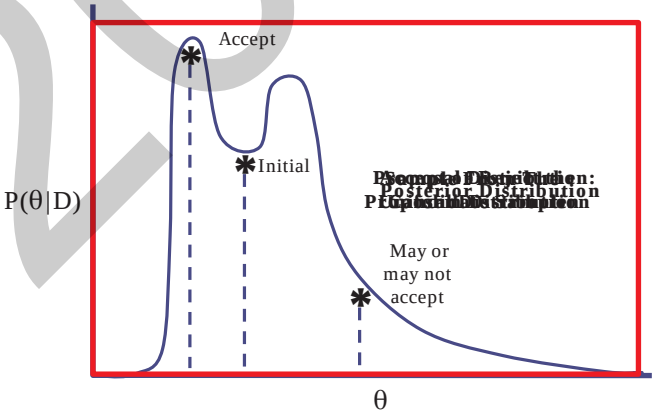
Correct Shape Approximation

$$\sum_{i=1}^{n}(y_i - f(x_i, \theta))^2 \leq \sum_{i=1}^{n}(y_i - f(x_i, \hat{\theta}))^2\left[1 + \frac{p}{n-p}F_{1-\alpha}(p, v)\right]$$

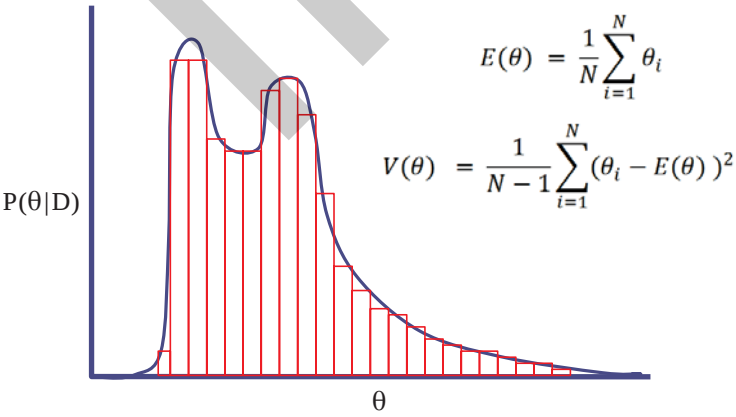
Correct Shape But Wrong Probability Content

Markov Chain Monte Carlo
Correct Shape and Correct Probability Content

Markov Chain Monte Carlo



Markov Chain Monte Carlo



$$E(\theta) = \frac{1}{N}\sum_{i=1}^{N}\theta_i$$

$$V(\theta) = \frac{1}{N-1}\sum_{i=1}^{N}(\theta_i - E(\theta))^2$$

Polymerization Case Study

•Estimate reactivity ratios and JCR in copolymer systems using Mayo-Lewis Model and Meyer-Lowry Model

•Free radical copolymerization of Di-*n*-Butyl Itaconate and Methyl Methacrylate

•Reactivity ratios are used for understanding the properties of polymeric materials

•Important in the control and operation of polymerization reactions

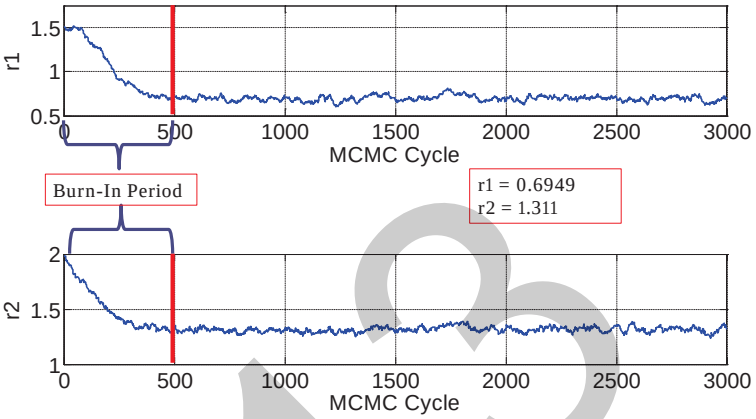
Mayo-Lewis Model

•Relationship between the instantaneous copolymer composition and the monomer feed composition

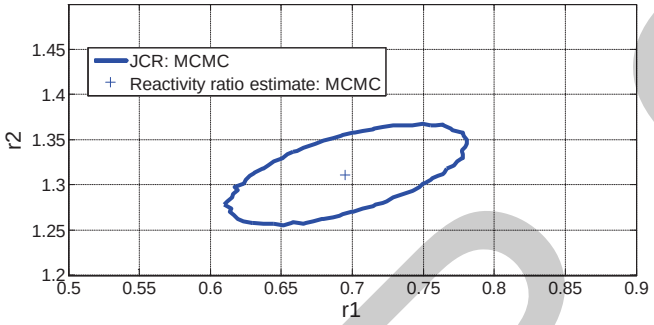
$$F_1 = \frac{r_1 f_1^2 + f_1 f_2}{r_1 f_1^2 + 2 f_1 f_2 + r_2 f_2^2}$$

•F_1 is the copolymer composition, f_1 is the monomer feed composition and r_1 and r_2 are the reactivity ratios

MCMC Output

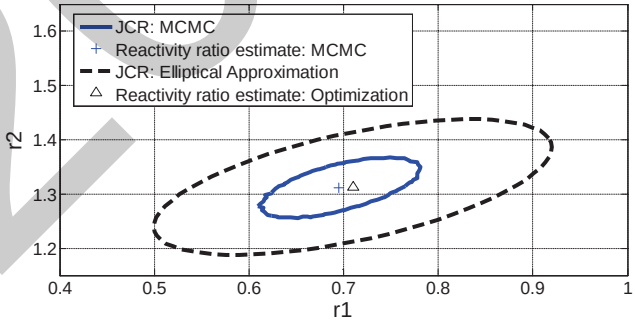


JCR for Mayo-Lewis Model



•Joint confidence region and point estimate obtained from MCMC simulation using Mayo-Lewis Model for the copolymerization of Di-*n*-Butyl Itaconate and Methyl Methacrylate

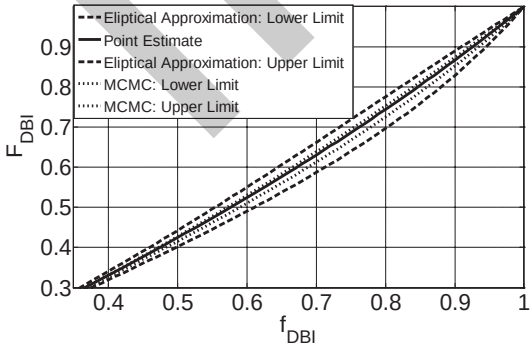
JCR for Mayo-Lewis Model



•The elliptical approach will yield JCRs with the wrong probability content

Kazemi et al., *Marcomol. React. Eng.* **2011,** *5*, 385-403

Mayo-Lewis Plot



•Different Mayo-Lewis curves obtained when using MCMC vs. Elliptical Approximation

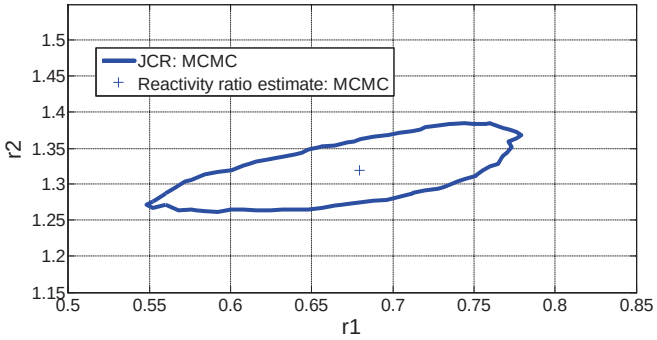
Meyer-Lowry Model

•Meyer-Lowry Model takes into account the conversion, X_N

$$X_N = 1 - \left(\frac{f_{10} - \bar{F}_1 X_n}{f_{10}(1 - X_n)}\right)^{\alpha} * \left(\frac{1 - X_n - f_{10} - \bar{F}_1 X_n}{(1 - f_{10})(1 - X_n)}\right)^{\beta} * \left(\frac{(\delta - f_{10})(1 - X_n)}{\delta - \delta X_n - f_{10} + \bar{F}_1 X_n}\right)^{\gamma}$$

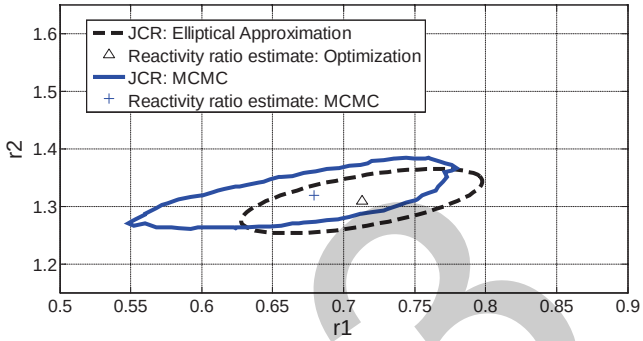
$$\alpha = \frac{r_2}{(1 - r_2)}, \quad \beta = \frac{r_1}{(1 - r_1)}, \quad \gamma = \frac{1 - r_1 r_2}{(1 - r_1)(1 - r_2)} \quad \delta = \frac{(1 - r_1)}{(2 - r_1 - r_2)}$$

JCR for Meyer-Lowry Model



•Joint confidence region and point estimate obtained from MCMC simulation using Meyer-Lowry Model for the copolymerization of Di-*n*-Butyl Itaconate and Methyl Methacrylate

JCR for Meyer-Lowry Model



•The elliptical approach will yield JCRs with the incorrect shape and probability content

Kazemi et al., *Marcomol. React. Eng.* **2011,** *5*, 385-403

Conclusions

•Applied MCMC methods to a polymerization case study

•Used a EVM model that can be used as a general regression framework

•Advantages of using MCMC

1. Easier approach when compared to optimization
2. Mayo-Lewis results showed that using a elliptical approach produces JCRs with the wrong probability content
3. Meyer-Lowry results showed that the elliptical approach does not provide the correct shape

Future Work

•WAT-Poly Program

- •Continue parameter estimation in more a complicated model

- •Previous work showed that accurate point estimates and JCR could not be obtained due to a complex sum of squares surface

•Computationally intensive

- • Use SHARCNET computer cluster to run program in parallel

Acknowledgements

•Professor Tom Duever

•Professor Park Reilly

•Professor Alexander Penlidis

•Niousha Kazemi

•Samira Masoumi and Yuncheng Du

•Funding from OGS and NSERC

Thank You for Listening

Metropolis Hastings Algorithm

Initial Guess X_o

↓

Sample point Y_t from a Proposal Distribution

↓

Acceptance Criterion

$$U \leq \min\left(1, \frac{\pi(Y)}{\pi(X)}\right)$$

Where U is a sample from a uniform distribution (0,1)

↓

If Candidate is Accepted: $X_{t+1} = Y_t$
Otherwise: $X_{t+1} = X_t$

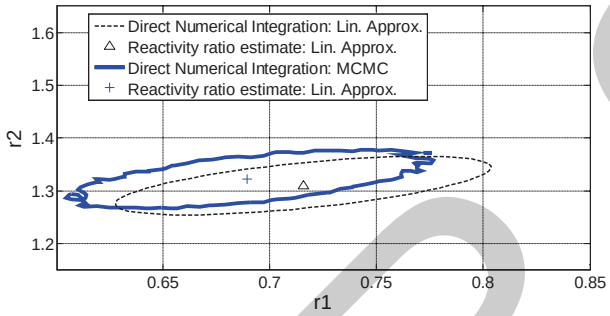
Integrated Cumulative Model

$$\frac{df_1}{dX_n} = \frac{f_1 - F_1}{1 - X_n} \qquad F_1 = \frac{r_1 f_1^2 + f_1 f_2}{r_1 f_1^2 + 2 f_1 f_2 + r_2 f_2^2}$$

$$F_1 = \frac{f_{10} - f_1(1 - X_n)}{X_n}$$

- The differential equation must be solved numerically along with the MCMC algorithm

JCR for Integrated Cumulative Model



The linear approach will again give JCRs that have incorrect probability content