

**Bayesian Learning and Inference  
in Recurrent Switching Linear Dynamical Systems**

Charbel Abi Younes, Marvyn Bailly,  
Bart Boom, Rohin Gilman, Daran Xu

May 1, 2023

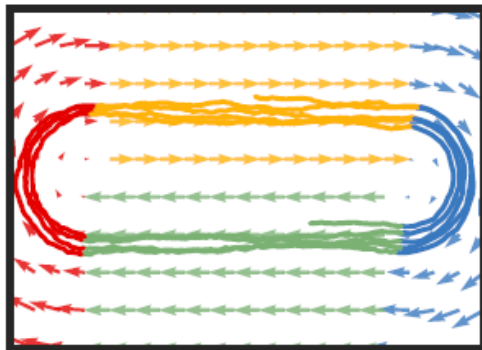
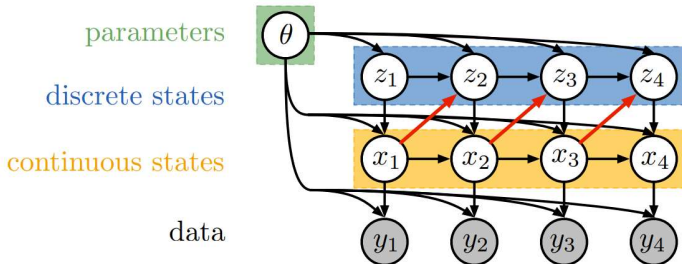


Figure: NASCAR Dynamical System

## SLDS and rSLDS

- Continuous latent state  $x_{t+1} = A_{z_{t+1}}x_t + b_{z_{t+1}} + \nu_t$ ,  $\nu_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, Q_{z_{t+1}})$
- Observation  $y_t = Cx_t + d + w_t$ ,  $w_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, S)$
- Discrete latent state  $z_t \in \{1, \dots, K\}$ 
  - SLDS [1]  $z_{t+1}|z_t \sim \pi_{z_t}$
  - rSLDS [2]  $z_{t+1}|z_t, x_t \sim \pi_{SB}(\nu_{t+1})$ ,  $\nu_{t+1} = R_{z_t}x_t + r_{z_t}$



## Stick Breaking Logistic Regression [3]

- $p(z|x) \sim \pi_{SB}(\nu)$ ,  $\nu = Rx + r$
  - Link function:  $\pi_{SB}(\nu) = \left( \pi_{SB}^{(1)}(\nu), \dots, \pi_{SB}^{(K)}(\nu) \right)$  (with  $\sigma(x) = \frac{e^x}{1+e^x}$ )
 
$$\pi_{SB}^{(k)}(\nu) = \begin{cases} \sigma(\nu_k) \prod_{j < k} \sigma(-\nu_j), & \text{if } k = 1, \dots, K-1 \\ \prod_{j=1}^K \sigma(-\nu_j), & \text{if } k = K \end{cases}$$
  - $p(z|x) \sim \prod_{k=1}^K \sigma(\nu_k)^{\mathbb{1}[z=k]} \sigma(-\nu_k)^{\mathbb{1}[z > k]}$  (Likelihood)
- With a Gaussian Prior  $p(x)$ , the posterior  $p(x|z) \propto p(x) * p(z|x)$  is non-Gaussian
  - Bayesian updating is not efficient
    - Gibbs Sampler: sampling  $x^{(i+1)} \sim p(x|z = z^{(i)})$ , sampling  $z^{(i+1)} \sim p(z|x = x^{(i+1)})$

## Polya-gamma augmentation [3]

$$\frac{(e^\nu)^a}{(1 + e^\nu)^b} = 2^{-b} \int_0^\infty e^{\kappa\nu} e^{-\omega\nu^2/2} p_{\text{PG}}(\omega \mid b, 0) d\omega \quad (\kappa = a - \frac{b}{2}) \quad (2.1)$$

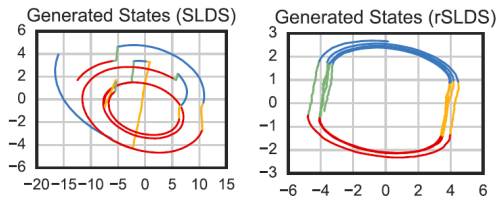
$$p(x_t | z_{t+1}) \propto \prod_{k=1}^{K-1} \frac{(e^{\nu_{t+1,k}})^{\mathbb{I}[z_{t+1}=k]}}{(1 + e^{\nu_{t+1,k}})^{\mathbb{I}[z_{t+1} \geq k]}} \quad (2.2)$$

$$p(x_t | z_{t+1}) = \int p(x_t, \omega_t | z_{t+1}) d\omega_t = \int p(x_t | z_{t+1}, \omega_t) p(\omega_t) d\omega_t \quad (2.3)$$

$$\omega_{t,k} \mid x_t, z_{t+1} \sim \text{PG}(\mathbb{I}[z_{t+1} \geq k], \nu_{t+1,k}) \quad (2.4)$$

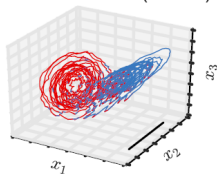
- $p(x_t | z_{t+1}, \omega_t)$  is Gaussian
- Thus instantiating these auxiliary variables in a Gibbs sampler enables efficient block updates

- SLDS vs rSLDS



- Canonical Dynamical System - Lorenz Attractor

Generated States (rSLDS)



- Van der Pol Oscillator
  - Canonical dynamical system with a stable limit cycle and bifurcations
- Double Pendulum
  - System with chaotic dynamics and multiple states (depending on the relative positions of the weights)



Guy A. Ackerson and King-Sun Fu.

On state estimation in switching environments.

*IEEE Transactions on Automatic Control*, 15(1):10–17, 1970.



David Barber.

Expectation correction for smoothed inference in switching linear dynamical systems.

*Journal of Machine Learning Research*, page 7(Nov):2515–2540, 2006.



Scott W Linderman, Matthew J Johnson, and Ryan P Adams.

Dependent multinomial models made easy: Stick-breaking with the poly-gamma augmentation.