# Online Variational Message Passing in Autoregressive Models



Albert Podusenko


Wouter M. Kouw


Bert de Vries


## Motivation

Autoregressive (AR) models are one of the most popular ways to describe different time-varying processes in nature, economics, etc. However, their parameters are often estimated in a batch manner which makes them inefficient for handling large-scale real-time data. In our work, we investigate the feasibility of online parameter estimation for these types of models. We translate the AR model to a probabilistic factor graph which takes advantage of the factorization of the model by implementing inference as a message passing algorithm. Due to the intractability of exact parameter inference for these types of models, sum-product message passing becomes impractical. This suggests to use alternative message passing algorithms based on approximate inference, e.g., variational message passing (VMP) which tries to find variational distributions that serve as good proxies for the exact solution. With VMP, the computations for online state and parameter estimation can be automated.

## Autoregressive State-Space Model

$$x_t = \sum_i^p \theta_i x_{t-i} + p_t$$
$$y_t = x_t + m_t$$

$$\begin{bmatrix} x_t \\ \vdots \\ x_{t-p+1} \end{bmatrix} = \begin{bmatrix} \Theta^{\mathrm{T}} & \\ \mathbf{I}_{p-1} & \mathbf{0} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ \vdots \\ x_{t-p} \end{bmatrix} + \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} p_t$$

$$y_t = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}^T \begin{bmatrix} x_t \\ \vdots \\ x_{t-p+1} \end{bmatrix} + m_t$$

$p_t \sim N(0, \gamma_x^{-1})$
$m_t \sim N(0, \gamma_y^{-1})$

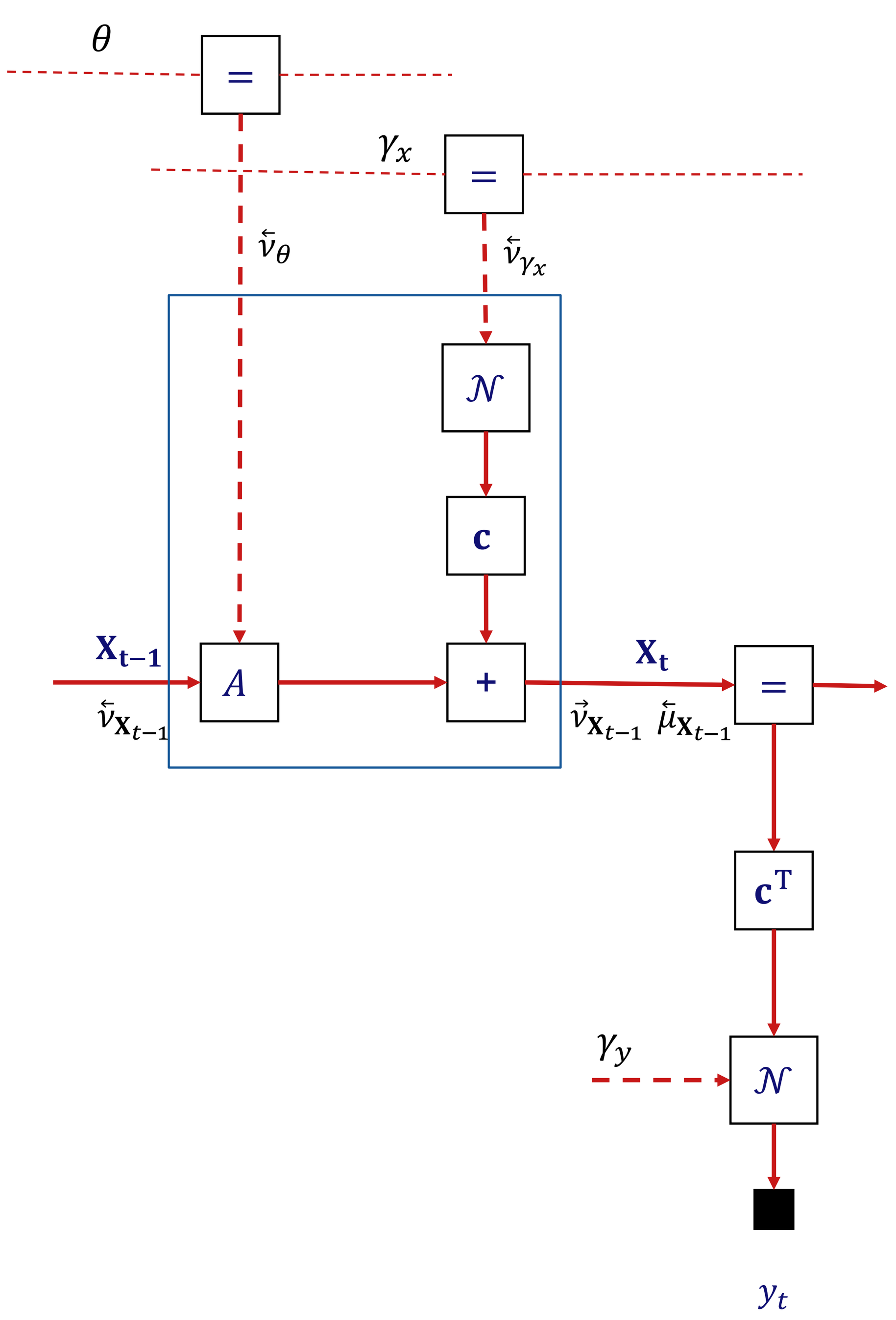
$$\mathbf{x_t} = \mathrm{A}(\Theta)\mathbf{x}_{t-1} + \mathbf{c}\mathrm{p_t}$$
$$y_t = \mathbf{c}^{\mathbf{T}}\mathbf{x}_t + m_t$$

## Probabilistic model

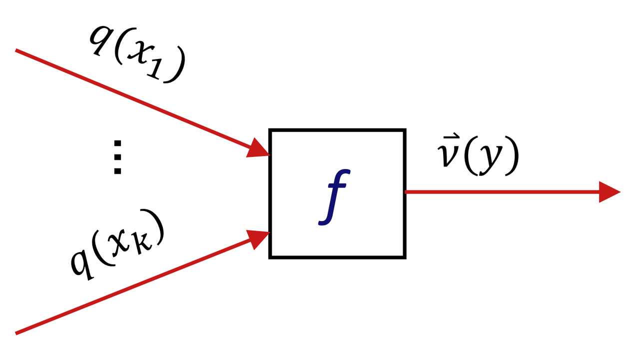
$$p(\mathbf{x}, \mathbf{y}, \Theta, \gamma_x, \gamma_y) = \underbrace{p(\mathbf{x_0})p(\Theta)p(\gamma_x)p(\gamma_y)}_{\text{priors}} \prod_{t=1}^{T} \underbrace{p(\mathbf{x}_t | \Theta, \mathbf{x}_{t-1}, \gamma_x)}_{\text{transition}} \underbrace{p(y_t | \mathbf{x}_{t-1}, \gamma_y)}_{\text{observation}}$$

*we are interested in the posterior:* $p(\boldsymbol{x}|\mathbf{y}_{1:t}) = \dfrac{\iint p(\mathbf{x}, \mathbf{y}, \Theta, \gamma_x, \gamma_y)\, \mathrm{d}\Theta \mathrm{d}\gamma_x}{p(y_t | y_{1:t-1})}$

## AR FFG



## VMP



**VMP** message:

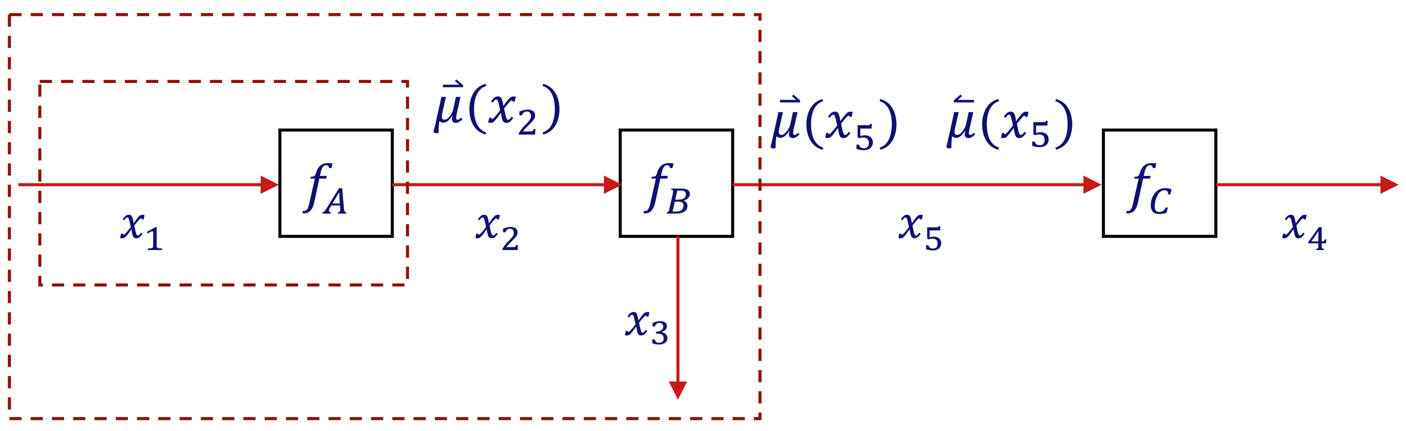
$$\vec{\nu}(y) \propto \exp\left(\int q(\mathbf{x}) \log f(y, x_1, \dots, x_k) d\mathbf{x}\right)$$

**Marginal** for $y$:

$$q(y) = \vec{\nu}(y)\overleftarrow{\nu}(y)$$

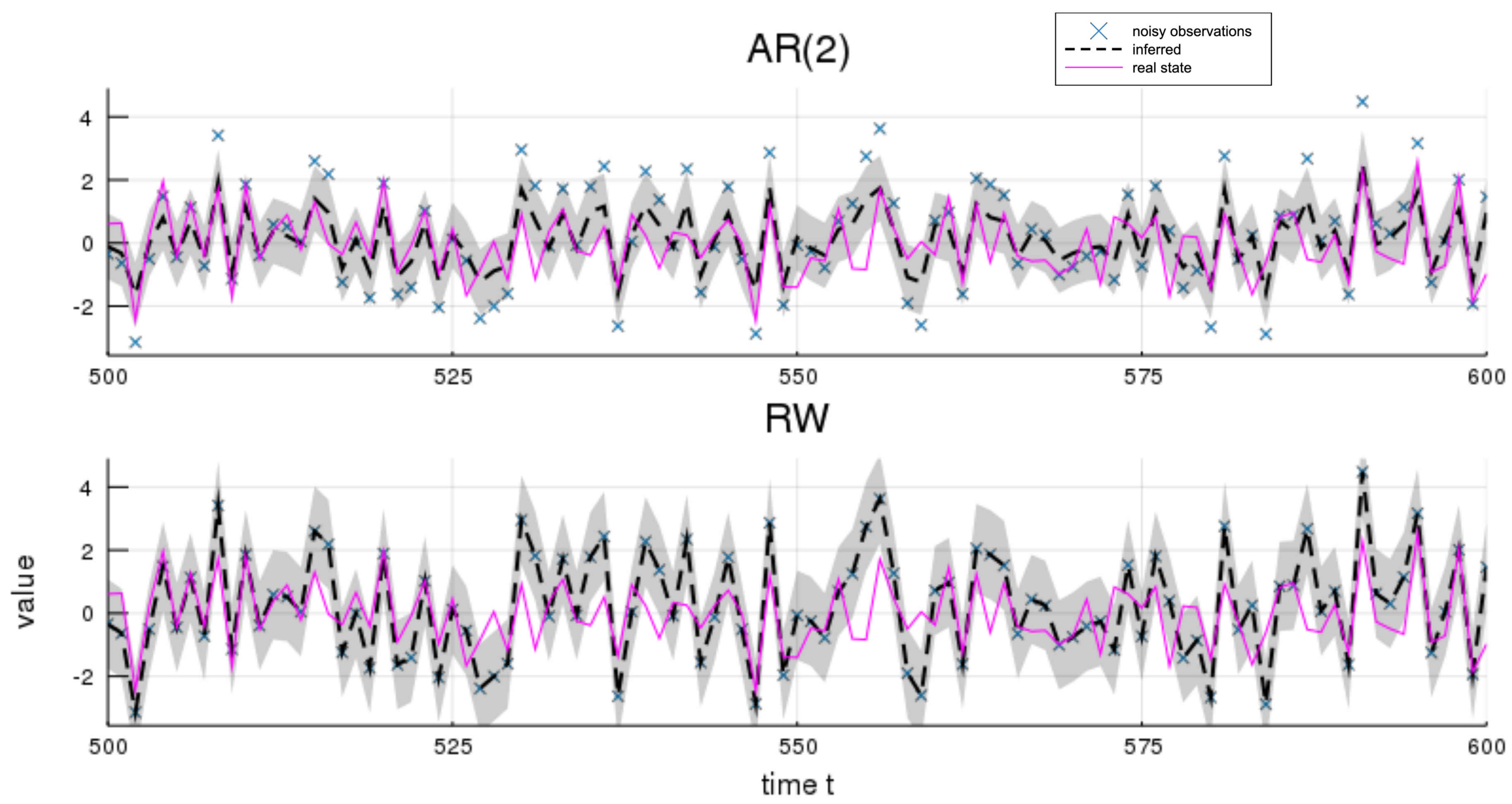
## Forney-style Factor Graphs

$$p(x_1, x_2, x_3, x_4, x_5) = f_A(x_1, x_2) f_B(x_2, x_3, x_5) f_C(x_4, x_5)$$



$$p(x_5) = \underbrace{\int \int f_B(x_2, x_3, x_5) \mathrm{d}x_3 \underbrace{\int f_A(x_1, x_2) \mathrm{d}x_1}_{\vec{\mu}(x_2)} \mathrm{d}x_2}_{\vec{\mu}(x_5)} \underbrace{\int f_C(x_4, x_5) \mathrm{d}x_4}_{\overleftarrow{\mu}(x_5)}$$

## Models comparison






https://github.com/biaslab/ForneyLab.jl