Inferring community characteristics in labelled networks

IIB Project

Lawrence Tray

Ioannis Kontoyiannis

June 5, 2021

Overview

Introduction

Motivation

Preliminaries

The stochastic block model (SBM)

Initial parameters:

- N – number of vertices
- B – number of blocks

SBM parameters:

- b – block membership vector
- e – block connectivity matrix
- k – degree sequence

$$A \sim \text{DC-SBM}_{\text{MC}}(b, e, k) \qquad (1)$$

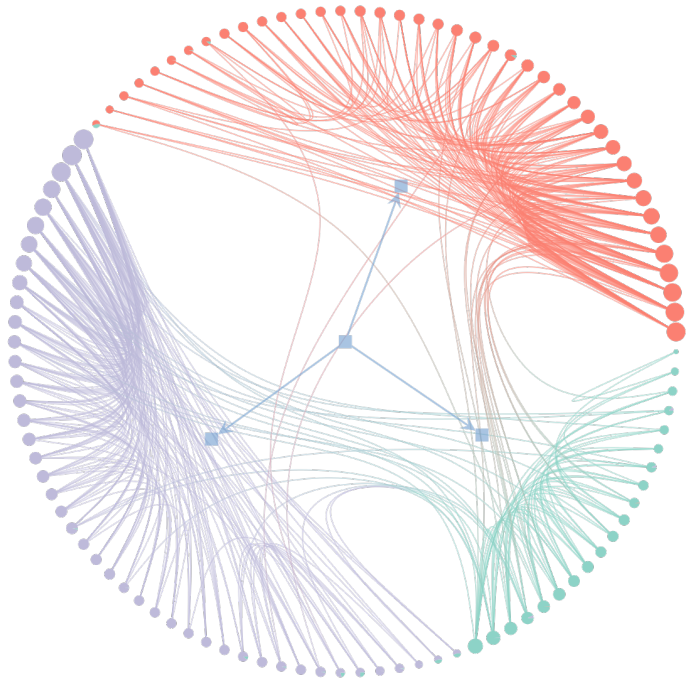

Figure: Typical SBM

The feature-first block model

The feature-first block model (FFBM)

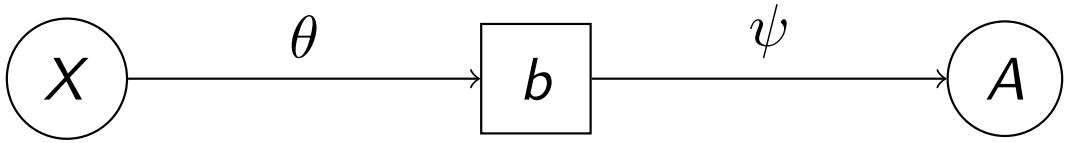

Figure: The feature-first block model (FFBM)

$$p(b|X;\theta) = \prod_{i\in[N]} \phi_{b_i}(x_i;\theta) = \prod_{i\in[N]} \frac{\exp(w_{b_i}^T x_i)}{\sum_{k\in[B]} \exp(w_k^T x_i)} \tag{2}$$

$$p(A|b;\psi) \sim \text{DC-SBM}_{\text{MC}}(b, \psi_e, \psi_k) \tag{3}$$

Inference

Inference procedure

We want to draw:

$$\theta^{(t)} \sim p(\theta|A, X). \tag{4}$$

We achieve this by:

$$b^{(t)} \sim p\Big(b|A, X\Big) \tag{5}$$

$$\theta^{(t)} \sim p\Big(\theta|X, b^{(t)}\Big) \tag{6}$$

Sampling sequence

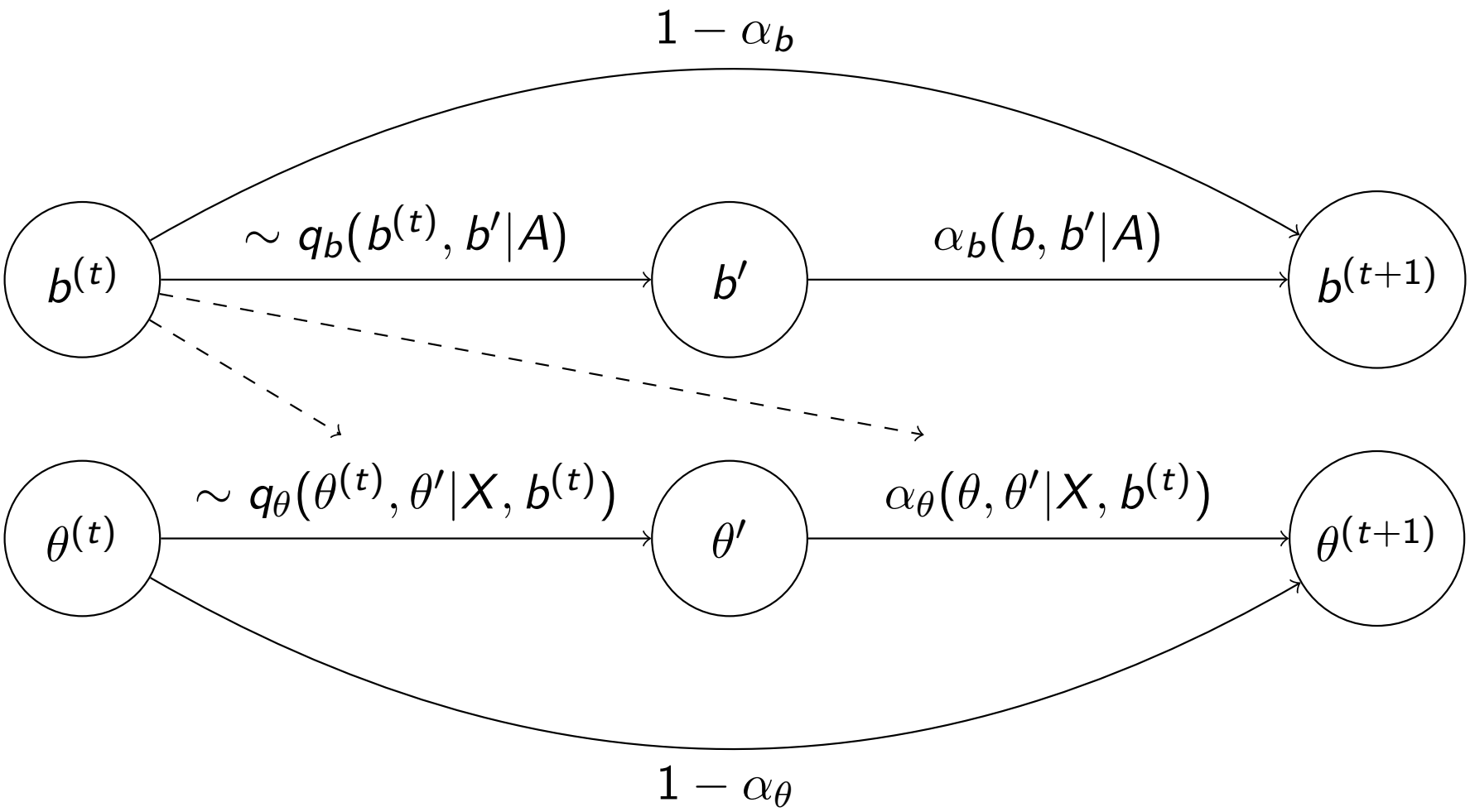

Figure: Sampling sequence.

Speed up computation

content...

Dimensionality Reduction

Experiments

Political Books

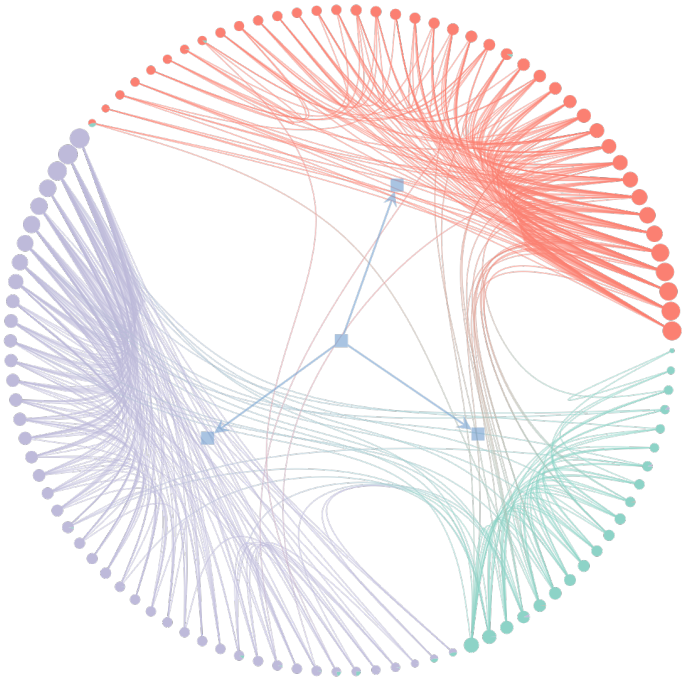

Figure: Polbooks

References