Project Presentation

CSC 696H - Advanced Topics in Probabilistic Graphical Models

Yang Hong
hong1@arizona.edu

University of Arizona

December 5, 2022

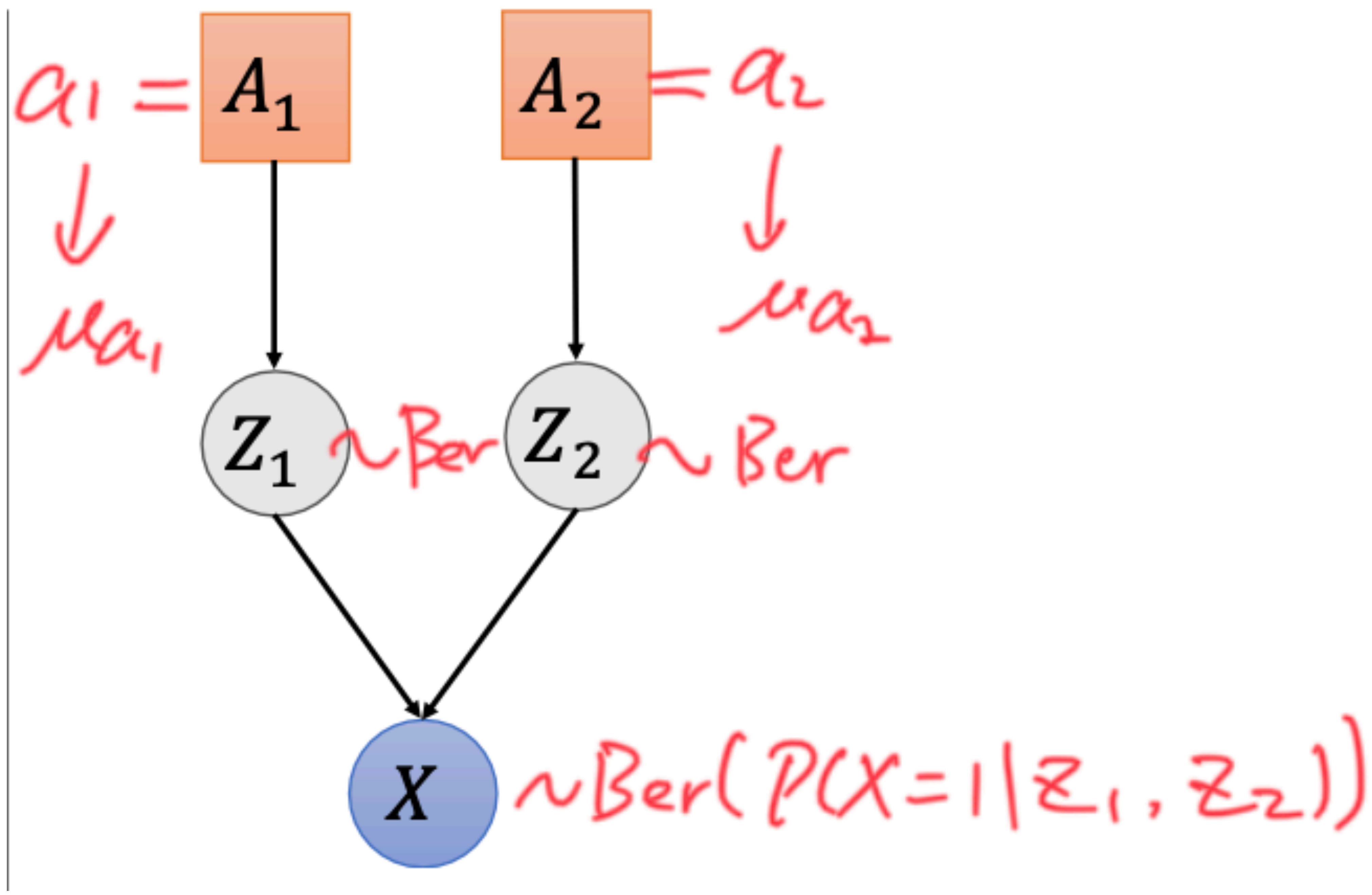
Problem Statement

- *Stochastic Rank-1 Bandits*
  - The problem of finding the maximum entry of a rank-1 matrix from noisy and adaptively-chosen observations
  - Motivation: a position-based click model
  - a model of people clicking on a list of K items out of L positions
  - *item* - ***attraction,*** each *position* - ***examination***: i.i.d. Bernoulli r.v.s
  - The *item* is *clicked* only if it is attractive and its position is examined
  - the pair of *item* and *position* that maximizes the probability of clicking is the maximum entry of a rank-1 matrix (outer product)
- *Graphical Models Meet Bandits: A Variational Thompson Sampling Approach*
  - influence diagram bandit framework -> rank-1 bandit
  - Thompson sampling approach with mean-field variational inference

Problem Statement

Algorithm 1 idTSvi: Influence diagram TS with variational inference.

Input: $\epsilon > 0$
Randomly initialize q
for $t = 1, \ldots, n$ **do**
  Sample θ_t proportionally to $q(\theta_t)$
  Take action $a_t = \arg\max_{a \in \mathcal{A}^K} r(a, \theta_t)$
  Observes x_t and receive reward $r(x_t, z_t)$
  Randomly initialize q
  Calculate $\mathcal{L}(q)$ using (3) and set $\mathcal{L}'(q) = -\infty$
  while $\mathcal{L}(q) - \mathcal{L}'(q) \geq \epsilon$ **do**
    Set $\mathcal{L}'(q) = \mathcal{L}(q)$
    for $\ell = 1, \ldots, t$ **do**
      Update $q_\ell(z_\ell)$ using (4), for all z_ℓ E
    end for
    Update $q(\theta)$ using (5) M
    Update $\mathcal{L}(q)$ using (3) maximize
  end while
end for

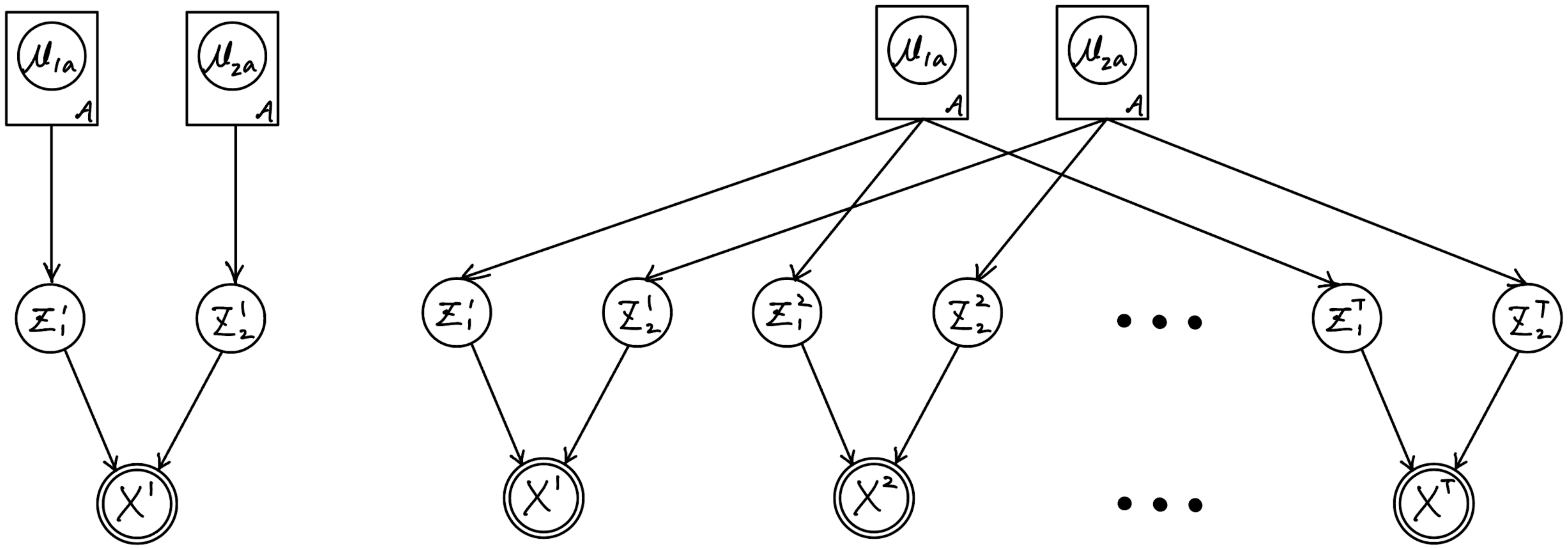


- Problems
  - idTSvi: E-, M-steps alternated until convergence, guaranteed
  - this approach is computationally expensive
  - idTSinc: an incremental variant of the algorithm in order to reduce the computational complexity

Problem Statement

- Extend existing Thompson sampling approach to the rank-1 bandit setting
- Expectation Propagation (EP) to approximate the true posterior
    - Regret lower bound is already optimal, want to achieve the same
    - If our approach is computationally more efficient
- a finite time T
- a total number of actions A
- parametrized by a model parameter $\theta = [\alpha_{a\in\{1,\ldots,A\}}, \beta_{a\in\{1,\ldots,A\}}]^{\mathrm{T}}$ where each action a is associated with θ_a
- replace decision nodes A_1, A_2 with Bernoulli r.v.s: μ_1, μ_2
- keep Z_1, Z_2, X the same

Approach



$$\mu_1 \sim Beta(\alpha_1, \beta_1)$$
$$\mu_2 \sim Beta(\alpha_2, \beta_2)$$
$$Z_1^t | \mu_1 \sim Ber(\mu_1)$$

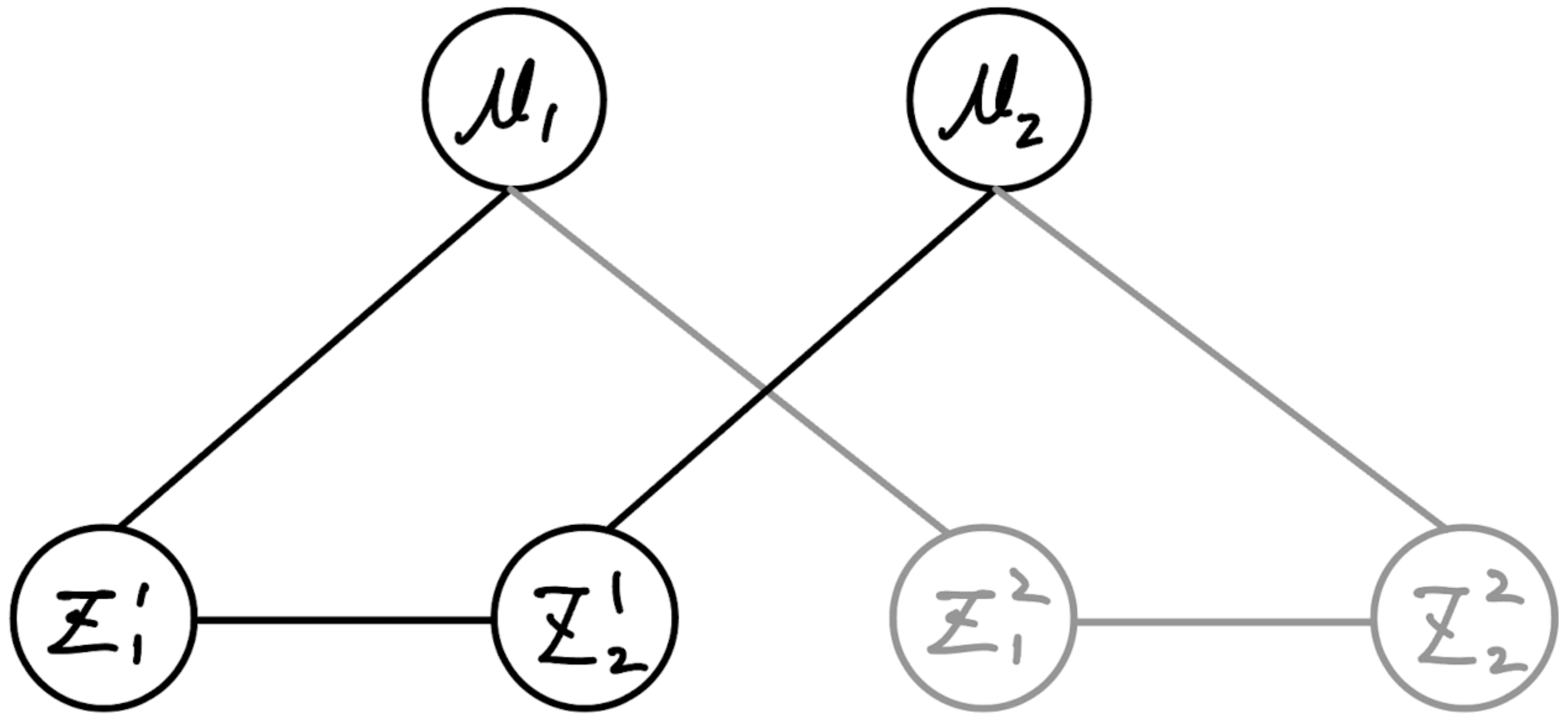
- $$Z_2^t | \mu_2 \sim Ber(\mu_2)$$

$$X^t | Z_1^t, Z_2^t \sim Ber(Z_1^t Z_2^t)$$

Approach

- instantaneous reward at round t
- $$r(a, \theta) = r(X^t, Z_1^t, Z_2^t \,|\, a, \theta) = X^t$$
- instantaneous regret at round t
- $$R(t \,|\, a, \theta) = r(a^*, \theta_*) - r(a^t, \theta_*)$$
- where $a^* = \text{argmax}_{a \in A}\ r(a, \theta_*)$, unique; θ_* is the true model parameter that is unknown to the learning agent
- cumulative regret
- $$R(T) = \sum_{t=1}^{T} R(t \,|\, a, \theta) = \sum_{t=1}^{T} \left(\ r(a^*, \theta_*) - r(a^t, \theta_*)\ \right)$$

Approach

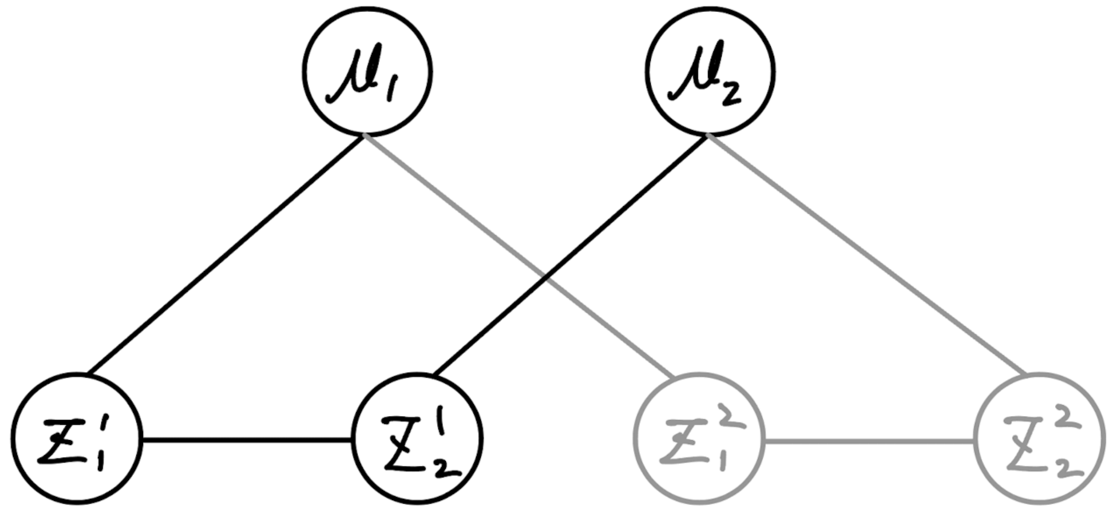


- convert directed PGM into pairwise MRF with nodes $\mathscr{V}$ and edges $\mathscr{E}$
- reduce the set of $\mu_{1a}, \mu_{2a}, a \in A$ into μ_1, μ_2 for simplicity
- target joint
- $$p(y) = \prod_{s,t \in \mathscr{E}} \psi_{s,t}(y_s, y_t)$$
- where $\psi_{s,t}(y_s, y_t)$ is the compatibility potential function between nodes y_s and y_t, and the true posterior $p(y;\theta) \propto p(y)$

Approach

- choose our approximating distribution $q(y;\theta)$ to be fully factorized
- $$q(y;\theta) = \prod_{s \in \mathcal{V}} q_s(y_s)$$
- where $q_s(y_s)$ is the approximating marginal
- $$q_s(y_s) \propto \prod_{t \in \Gamma(s)} m_{t \to s}(y_s)$$
- where $\Gamma(s)$ is the set of nodes neighboring s
- fully factorized $q(y;\theta)$ implies that the exponential approx. $m_{s,t}(y_s, y_t)$ to the compatibility potentials $\psi_{s,t}(y_s, y_t)$ also have a factorized form
- $$m_{s,t}(y_s, y_t) = m_{s \to t}(y_t) \cdot m_{t \to s}(y_s)$$

Approach



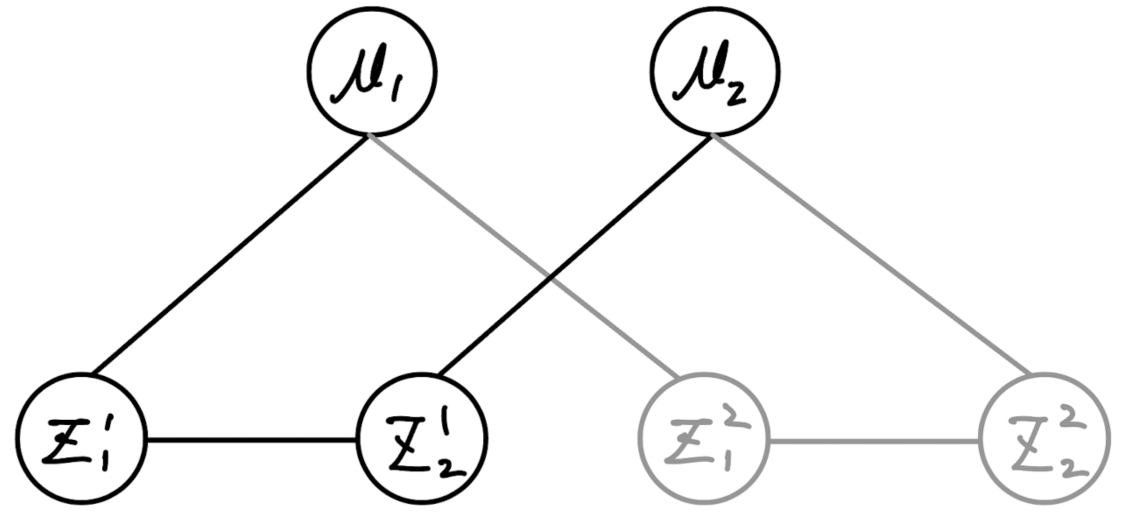
- we show updates for marginal $q(z_1), q(\mu_1)$
- same procedure applies to all
- all updates are for current time t, thus omit the superscript t for clarity
- shorten $m_{s \to t}(y_t)$ to $m_s(y_t)$ for simplicity
- **Step 1**: choose message $m_{1,2}(z_1, z_2)$ to refine. Remove its effects from current $q(y; \theta)$ by

$$q_{1\backslash 2}(z_1) \propto \frac{q_1(z_1)}{m_{z_2}(z_1)} = \frac{m_{\mu_1}(z_1) m_{z_2}(z_1)}{m_{z_2}(z_1)} = m_{\mu_1}(z_1)$$

- $$q_{2\backslash 1}(z_2) \propto \frac{q_2(z_2)}{m_{z_1}(z_2)} = \frac{m_{\mu_2}(z_2) m_{z_1}(z_2)}{m_{z_1}(z_2)} = m_{\mu_2}(z_2)$$

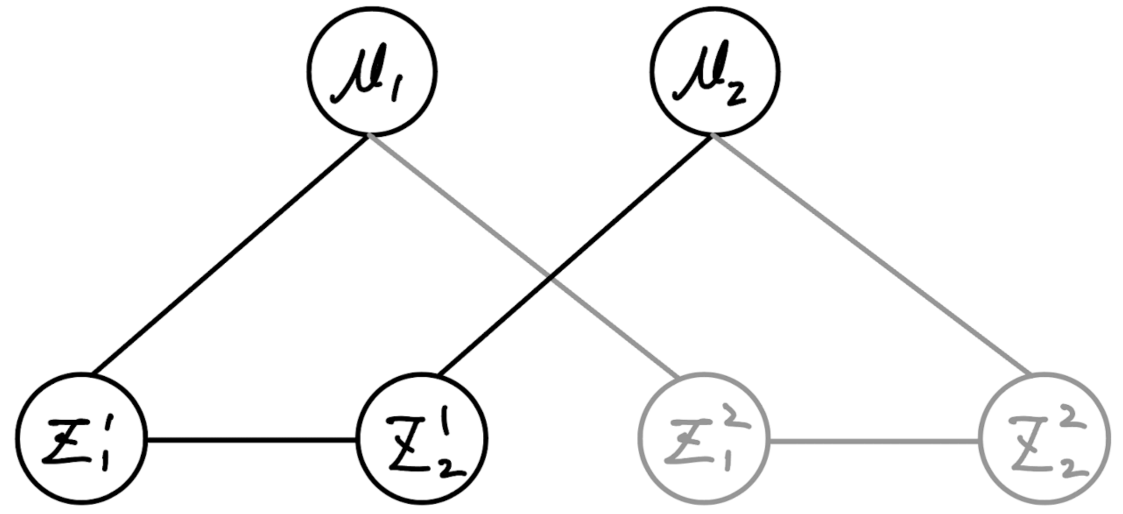
- where $m_{\mu_1}(z_1) \propto Ber(z_1 \,|\, \pi_{\mu_1})$ and $m_{\mu_2}(z_2) \propto Ber(z_2 \,|\, \pi_{\mu_2})$

Approach



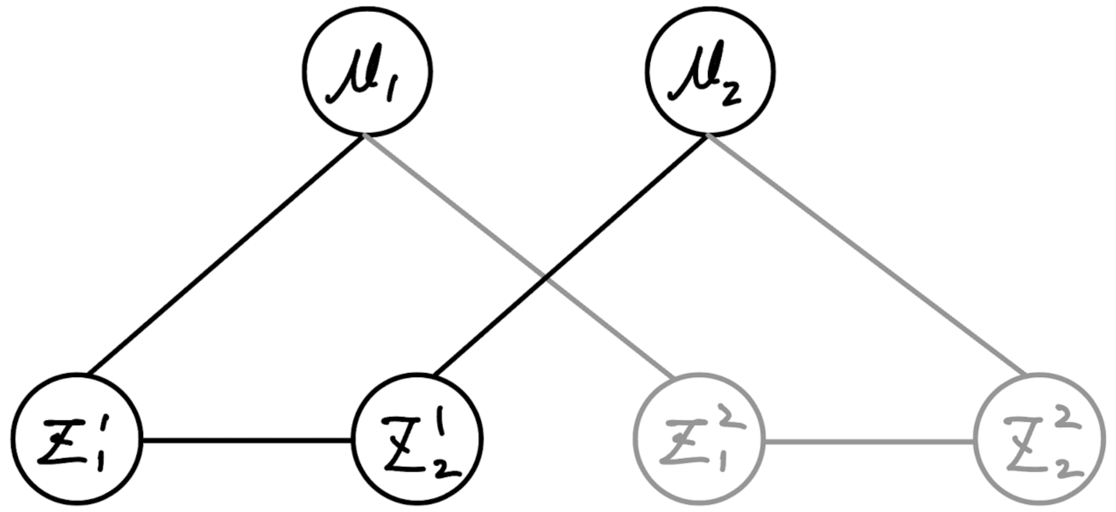
- **Step 2**: derive the local approximation to true posterior
- define the local approximation to true joint
- $$\hat{p}(z_1, z_2) \propto m_{\mu_1}(z_1) m_{\mu_2}(z_2) \psi_{1,2}(z_1, z_2)$$
  $$\propto Ber(z_1 \mid \pi_{\mu_1}) Ber(z_2 \mid \pi_{\mu_2}) Ber(x_1 \mid z_1 \cdot z_2)$$
- obtain the marginal of z_1
- $$\hat{p}(z_1) = \sum_{z_2} \hat{p}(z_1, z_2) = Ber(z_1 \mid \pi_{\mu_1})(1 - \pi_{\mu_2}) Ber(x_1 \mid 0) + Ber(z_1 \mid \pi_{\mu_1}) \pi_{\mu_2} Ber(x_1 \mid z_1)$$
- derive the local approximation to true posterior
- $$\hat{p}(z_1 \mid x_1) \propto \begin{cases} \pi_{\mu_1} \pi_{\mu_2} \mathbb{I}(z_1 = 1), \text{ if } x_1 = 1 \\ \begin{cases} \pi_{\mu_1}(1 - \pi_{\mu_2}) \mathbb{I}(z_1 = 1) \\ (1 - \pi_{\mu_1}) \mathbb{I}(z_1 = 0) \mathbb{I}(z_1 = 0) \end{cases}, \text{ if } x_1 = 0 \end{cases}$$

Approach



- **Step 3:** update approximating marginal $q_1^{new}(z_1)$
- this is equivalent to minimize the KL-divergence between the local approximation to true marginal $\hat{p}$ and the approximating distribution q
- $$q^{new} = \arg\min_q \; KL(\hat{p}\,||\,q)$$
- in our case
- $$q_1(z_1) \propto Ber(z_1\,|\,\pi_1) = \hat{p}(z_1\,|\,x_1)$$
- derive the parameter π_1^{new} update
- $$\pi_1^{new} = \mathbb{E}_{\hat{p}}\left[z_1\right] = \sum_{z_1=0}^{1} z_1 \cdot \hat{p}(z_1\,|\,x_1) = 0 \cdot \hat{p}(z_1=0\,|\,x_1) + 1 \cdot \hat{p}(z_1=1\,|\,x_1) = \hat{p}(z_1=1\,|\,x_1)$$
- $$\text{if } x_1 = 1,\ \pi_1^{new} \propto \pi_{\mu_1}\pi_{\mu_2}\mathbb{I}(z_1=1) = 1;$$ $$\text{if } x_1 = 0,\ \pi_1^{new} \propto \pi_{\mu_1}(1-\pi_{\mu_2})\mathbb{I}(z_1=1)\,.$$

Approach



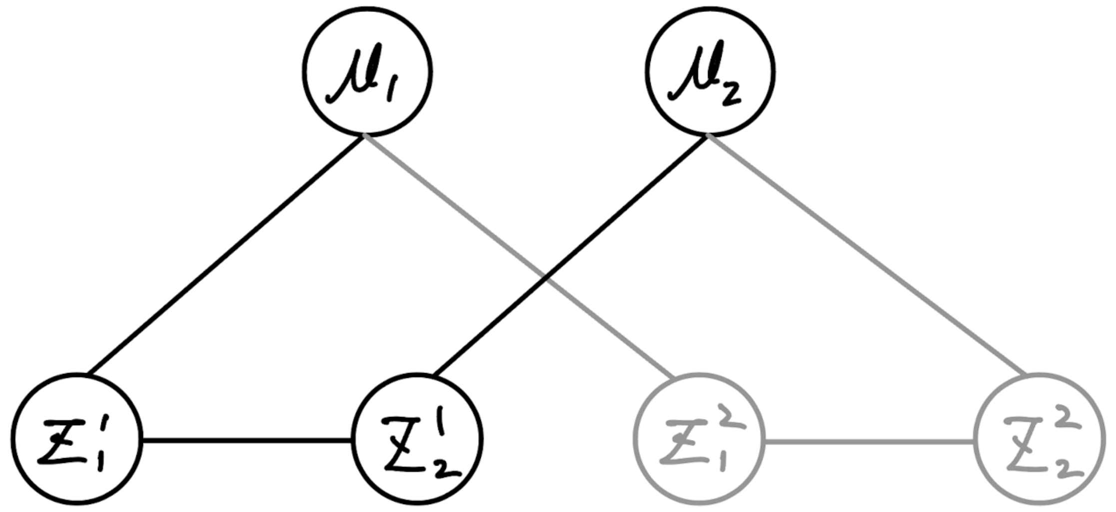
- **Step 4:** refine message $m_{1,2}(z_1, z_2)$ by

$$m_{z_2}^{new}(z_1) \propto \frac{q_1^{new}(z_1)}{q_{1\backslash 2}(z_1)} \propto \frac{q_1^{new}(z_1)}{m_{\mu_1}(z_1)} \propto \frac{Ber(z_1 \,|\, \pi_1^{new})}{Ber(z_1 \,|\, \pi_{\mu_1})} \propto Ber(z_1 \,|\, \frac{\pi_1^{new}}{\pi_{\mu_1}})$$

- $$m_{z_1}^{new}(z_2) \propto \frac{q_2^{new}(z_2)}{q_{2\backslash 1}(z_2)} \propto \frac{q_2^{new}(z_2)}{m_{\mu_2}(z_2)} \propto \frac{Ber(z_2 \,|\, \pi_2^{new})}{Ber(z_2 \,|\, \pi_{\mu_2})} \propto Ber(z_2 \,|\, \frac{\pi_2^{new}}{\pi_{\mu_2}})$$

- **Step 5**: refine finished, update approximating posterior $q(y;\theta)$ using the equation defined earlier $q(y;\theta) = \prod_{s \in \mathcal{V}} q_s(y_s)$

Approach

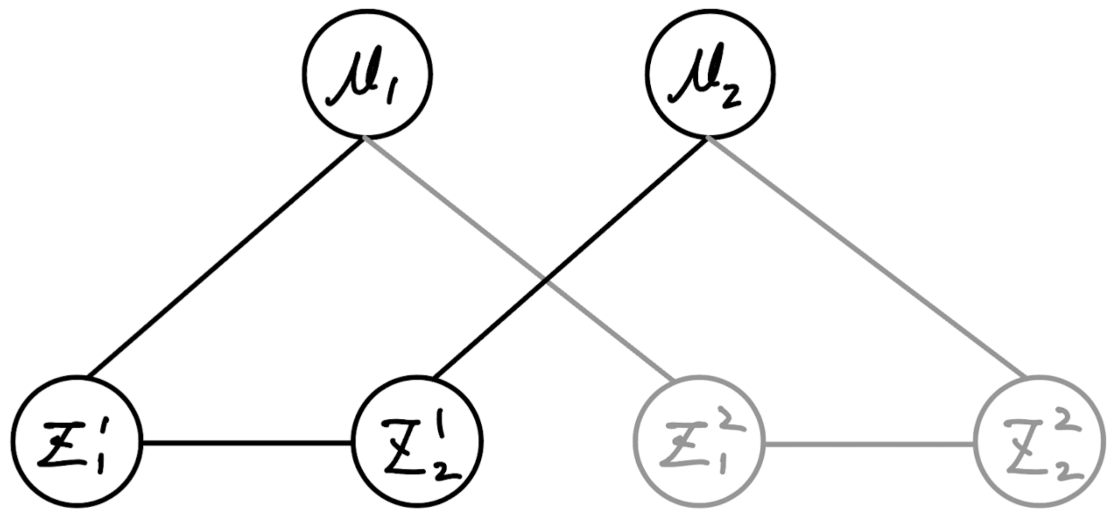


- Update for $q^{new}(\mu_1)$
- $$q^{new}(\mu_1) = \arg\min_q \; KL(\hat{p}(\mu_1)\,||\,q(\mu_1))$$
- where the local approximation to marginal of μ_1

$$\hat{p}(\mu_1) \propto \sum_{z_1^t} \hat{p}(z_1^t, \mu_1)$$

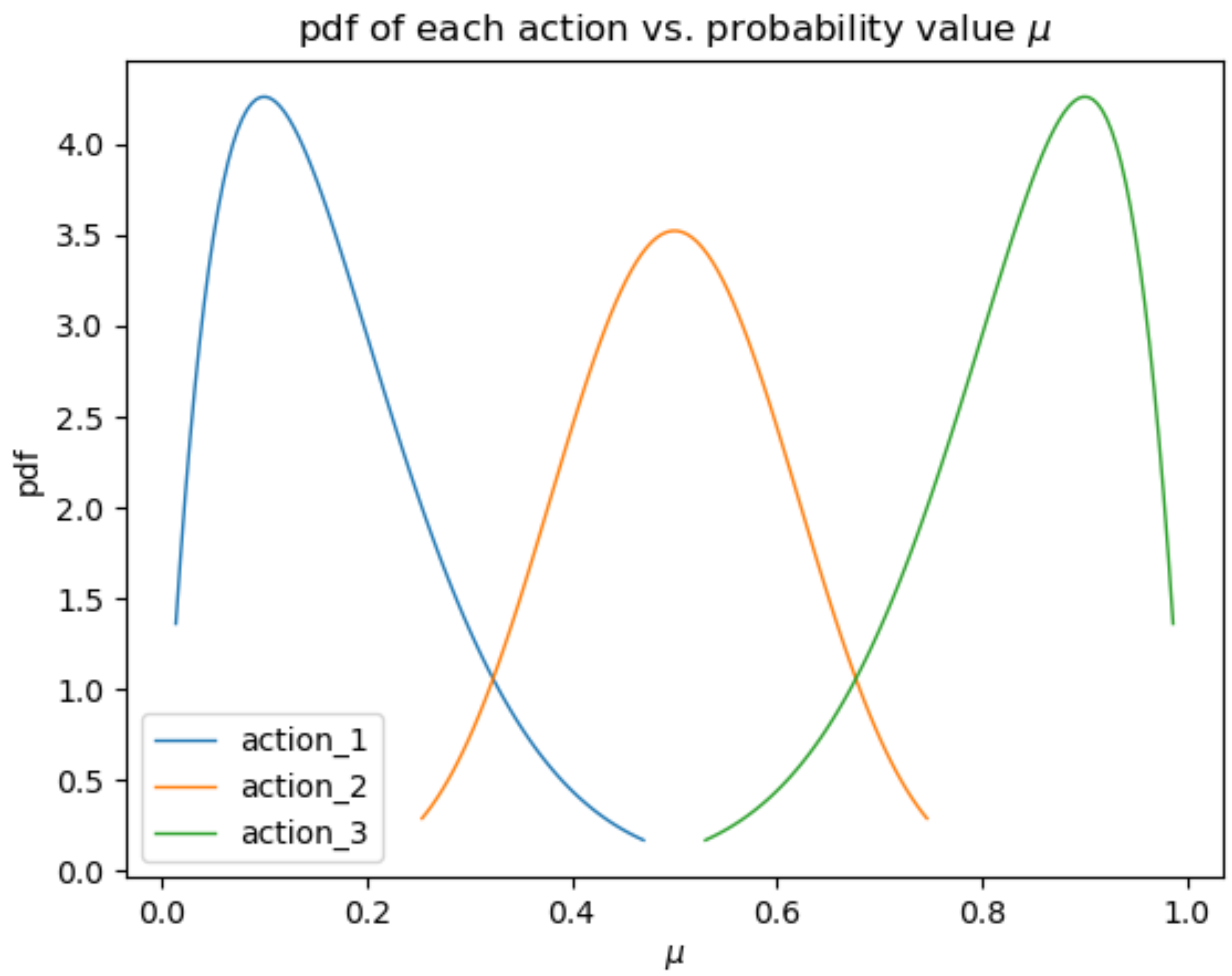
- $$\propto (1 - \frac{\pi_1^t \alpha_1'}{\alpha_1' + \beta_1'}) Beta(\mu_1 \,|\, \alpha_1', \beta_1' + 1) + \frac{\pi_1^t \alpha_1'}{\alpha_1' + \beta_1'} Beta(\mu_1 \,|\, \alpha_1' + 1, \beta_1')$$
- approx. $\hat{p}(\mu_1)$ by moment matching of Beta distribution's sufficient statistics
- $$\mathbb{E}_{\hat{p}}\left[\log(\mu_1)\right] = \psi(\alpha_1') - \psi(\alpha_1' + \beta_1')$$
  $$\mathbb{E}_{\hat{p}}\left[\log(1 - \mu_1)\right] = \psi(\beta_1') - \psi(\alpha_1' + \beta_1')$$
- where $\psi(\cdot)$ is the digamma function
- $$\psi(x) = \frac{d}{dx}\log\Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}$$

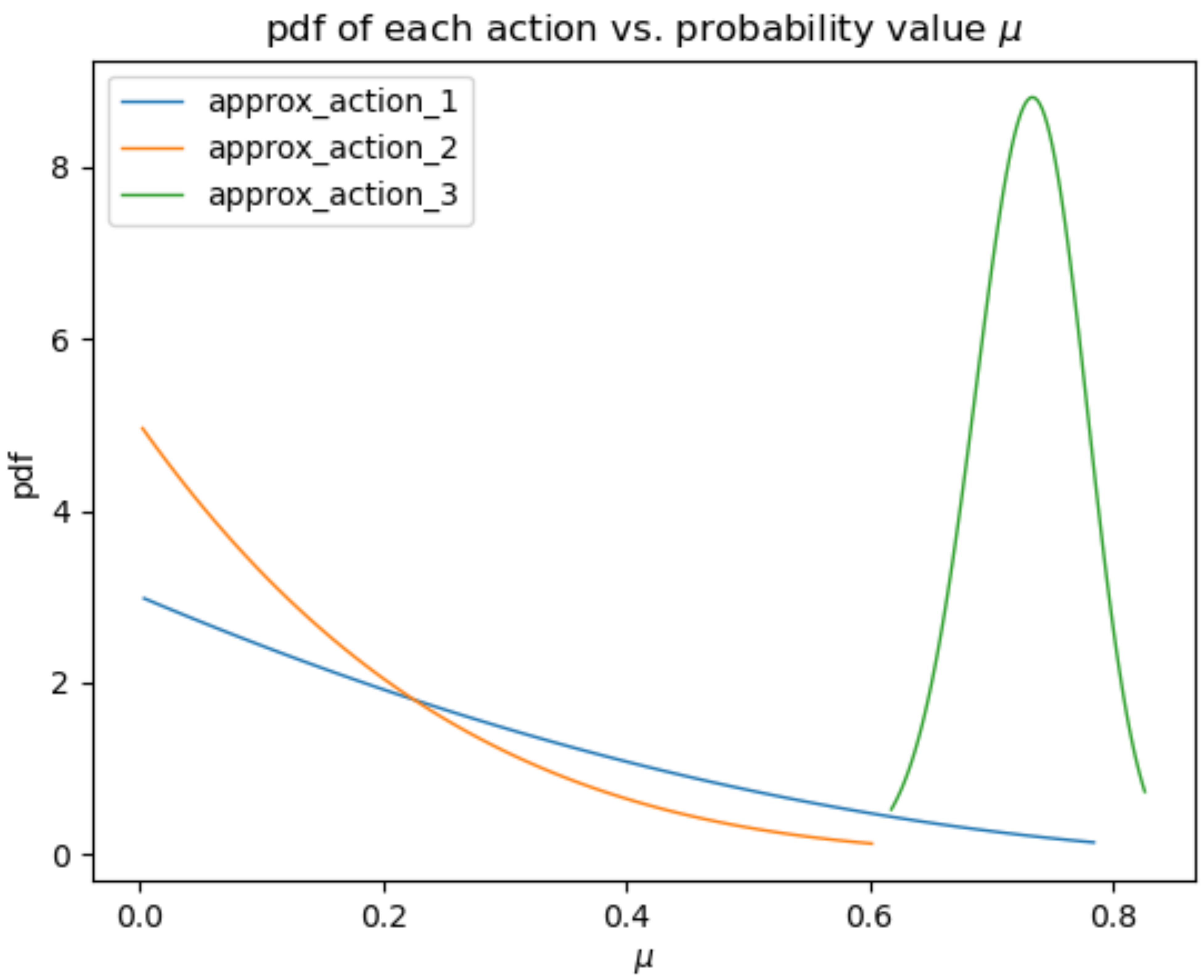
Approach



- Newton's method to solve for the root
- L-BFGS algorithm
  - which is in the family of quasi-Newton methods that finds a local minimum of an objective function
- Update for $\alpha_1^{new}, \beta_1^{new}$
- $\alpha_1^{new}, \beta_1^{new} = \arg\min_{\alpha_1', \beta_1'} \left[\frac{1}{2} \left[\mathbb{E}_{\hat{p}} \left[\log(\mu_1) \right] - \psi(\alpha_1') + \psi(\alpha_1' + \beta_1') \right]^2 + \frac{1}{2} \left[\mathbb{E}_{\hat{p}} \left[\log(1 - \mu_1) \right] - \psi(\beta_1') + \psi(\alpha_1' + \beta_1') \right]^2 \right]$

Preliminary result





Preliminary result