

XI. REFERENCE FOR MMSB

Among the few adaptations of HDP for networks, who generalized the MMSB, earlier work used the CRF to quantify degree of membership in small networks [17]. A more recent work [18] put the HDP in a dynamic framework to model the membership of node in networks consecutive times.

XII. REFERENCE FOR ILFM

The two seminal work for ILFM are [9] developed for dyadic relation. And a more general model BMF [7] for rectangular matrices. An adaptation of ILFM for non-negative weights in the interaction matrix has been done in [19]. The authors used Hamiltonian MCMC updates to learn those weights.

XIII. A RAJOUTER DANS BURSTINESS, LA SUITE DE LA PROPOSITION POUR LE BURSTINESS DANS LES FEATURES/CLASS

The models of interest draws their features from either the HDP or the IBP who are build using the conjugate property of exponential family, respectively Multinomial-Dirichlet and Bernoulli-Beta. The posterior distribution can be in both case collapsed, leading to the following gibbs updates:

$$\mathbf{p}(f_{ik} = 1 \mid F^{-ik}) \propto m^{-ik} \quad (79)$$

$$\mathbf{p}(z_{i \rightarrow j} = k \mid Z^{-ij}) \propto N_{ik}^{-ij} + \alpha_0 \quad (80)$$

The fact the collapse form of the posterior for a feature depend on the count of previously assigned features enable the burstiness behavior at the feature level.

XIV. NONPARAMETRIC UPDATE...

A. ILFM

The probability to have k_i^{new} features is proportional to this density:

$$\mathbf{p}(k_i^{new} \mid Y, F, \alpha) \propto \mathbf{p}(Y \mid F^{new}) \mathbf{p}(k_i^{new} \mid \alpha) \quad (81)$$

The probability of having k_i^{new} features is drawn from a $\text{Poisson}(\frac{\alpha}{N})$ distribution in the IPB process. However, we also need to sample the \mathbf{w}_i weights associated with those features, and in our case, the density of weights are not conjugate of the likelihood. In consequence, we cannot integrate them out as explicitly assumed in the equation (81). We follow the approach of [7] to jointly sample the new features and the weights using a Metropolis-Hasting method. Thanks to the exchangeability of the IPB we only need to consider features unique to entity i to either propose or reject new features. Therefore, we reference by a superscript B the set of model parameters corresponding to unique features. A convenient choice of jumping distribution is:

$$J(k_i^{new}, \phi_i^{new} \mid k_i^B, \phi_i^B) = \text{Poisson}(k_i^{new} \mid \alpha) \mathcal{N}(\phi_i^{new} \mid \sigma_w) \quad (82)$$

The acceptance ratio can thus be written as:

$$r_{B \rightarrow new} = \frac{\mathbf{p}(k_i^{new}, \phi_i^{new} \mid Y, \Phi, F, \alpha) J(k_i^B, \phi_i^B \mid k_i^{new}, \phi_i^{new})}{\mathbf{p}(k_i^B, \phi_i^B \mid Y, \Phi, F, \alpha) J(k_i^{new}, \phi_i^{new} \mid k_i^B, \phi_i^B)} \quad (83)$$

When replacing by equation (81) and (82), the acceptance ratio simplify to ratio of data likelihoods:

$$r_{B \rightarrow new} = \frac{\mathbf{p}(Y \mid F^{new}, \Phi^{new})}{\mathbf{p}(Y \mid F, \Phi)} \quad (84)$$

B. IMMSB

In the latent class model, the collapsed gibbs sampling approach allow us to only sample the classes couple for each observations:

$$\mathbf{p}(z_{ji} = k, z_{ij} = k' \mid \cdot) \propto \mathbf{p}(z_{ji} = k \mid \cdot) \mathbf{p}(z_{ij} = k' \mid \cdot) f_{(k,k')}^{-ji}(y_{ji}) \quad (85)$$

The class assignment updates for the node couple are:

$$\mathbf{p}(z_{i \rightarrow j} = k \mid Z^{-ij}) \propto N_{ik}^{-ij} + \alpha_0 \beta_k \quad (86)$$

$$\mathbf{p}(z_{i \leftarrow j} = k \mid Z^{-ij}) \propto N_{jk}^{-ij} + \alpha_0 \beta_k \quad (87)$$

Additionally, there is a non-null probability to create a new classes with probability:

$$\mathbf{p}(z_{i \rightarrow j} = k \mid Z^{-ij}) \propto \alpha_0 \beta_u \quad (88)$$

$$\mathbf{p}(z_{i \leftarrow j} = k \mid Z^{-ij}) \propto \alpha_0 \beta_u \quad (89)$$

$$(90)$$

Note that β as well as and the tables configuration in the CRF is sampled as proposed [6].

REFERENCES

- [1] Wray Buntine and Aleks Jakulin. Discrete component analysis. pages 1–33, 2006.
- [2] Edoardo M Airoldi, David Blei, Elena A Erosheva, and Stephen E Fienberg. Handbook of mixed membership models and their applications. 2014.
- [3] Konstantina Palla, David Knowles, and Zoubin Ghahramani. An infinite latent attribute model for network data. *arXiv preprint arXiv:1206.6416*, 2012.
- [4] Prem K Gopalan and David M Blei. Efficient discovery of overlapping communities in massive networks. *Proceedings of the National Academy of Sciences*, 110(36):14534–14539, 2013.
- [5] Edoardo M Airoldi, David M Blei, Stephen E Fienberg, and Eric P Xing. Mixed membership stochastic blockmodels. pages 33–40, 2009.
- [6] Yee Whye Teh, Michael I Jordan, Matthew J Beal, and David M Blei. Hierarchical dirichlet processes. *Journal of the american statistical association*, 101(476), 2006.
- [7] Edward Meeds, Zoubin Ghahramani, Radford M Neal, and Sam T Roweis. Modeling dyadic data with binary latent factors. pages 977–984, 2006.
- [8] Thomas L Griffiths and Zoubin Ghahramani. The indian buffet process: An introduction and review. *The Journal of Machine Learning Research*, 12:1185–1224, 2011.
- [9] Kurt Miller, Michael I Jordan, and Thomas L Griffiths. Nonparametric latent feature models for link prediction. pages 1276–1284, 2009.
- [10] Miller McPherson, Lynn Smith-Lovin, and James M Cook. Birds of a feather: Homophily in social networks. *Annual review of sociology*, pages 415–444, 2001.
- [11] Paul F Lazarsfeld, Robert K Merton, et al. Friendship as a social process: A substantive and methodological analysis. *Freedom and control in modern society*, 18(1):18–66, 1954.

- [12] Timothy La Fond and Jennifer Neville. Randomization tests for distinguishing social influence and homophily effects. In *Proceedings of the 19th international conference on World wide web*, pages 601–610. ACM, 2010.
- [13] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.
- [14] Kenneth W Church and William A Gale. Poisson mixtures. *Natural Language Engineering*, 1(02):163–190, 1995.
- [15] Stéphane Clinchant and Eric Gaussier. The bnb distribution for text modeling. In *Advances in Information Retrieval*, pages 150–161. Springer, 2008.
- [16] Stéphane Clinchant and Eric Gaussier. Information-based models for ad hoc ir. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, pages 234–241. ACM, 2010.
- [17] Phaedon-Stelios Koutsourelakis and Tina Eliassi-Rad. Finding mixed-memberships in social networks. pages 48–53, 2008.
- [18] Xuhui Fan, Longbing Cao, RY Xu, and Richard Yi. Dynamic infinite mixed-membership stochastic blockmodel. *arXiv preprint arXiv:1306.2999*, 2013.
- [19] Morten Mørup, Mikkel N Schmidt, and Lars Kai Hansen. Infinite multiple membership relational modeling for complex networks. pages 1–6, 2011.