Batched Adaptive Network Formation

Preliminary Working Draft

Yan Xu* and Bo Zhou†

February 2025

Abstract

Networks play a crucial role in many real-world applications, such as team formation in workplaces and connection recommendations on social platforms, among others. This paper proposes an adaptive, or *online*, policy for decision-makers who regulate network formation to maximize overall output. Our approach builds on the weighted stochastic block model (WSBM), where nodes represent agents, edges represent agent pairs, and edge weights capture their pairwise outcomes. WSBM captures agents' heterogeneous, unobservable characteristics through latent types, which determine their complementarities. Furthermore, it imposes no functional form restrictions on these complementarities, allowing for flexible modeling of agent interactions.

We frame the network formation policy as a *batched multi-armed bandit* problem, balancing the trade-off between exploration (learning agents' latent types and their complementarities) and exploitation (forming effective networks). We establish a *batched local asymptotic normality* result and show the asymptotic equivalence between the maximum likelihood estimates and those obtained from a variational approximation of the intractable likelihood. These results allow us to treat batched variational estimates as normal signals, facilitating Bayesian updating to obtain posterior distributions based on completed batches, which are then incorporated into a maximum-weight matching problem to determine the policy for the next batch. Simulation results show that our algorithm effectively improves outcomes within a few batches while refining parameter estimates with increasing accuracy. Moreover, it remains effective in nonstationary settings where the agent pool evolves over time, reflecting real-world dynamics.

Keywords: Adaptive network formation, algorithm design, weighted stochastic block model, batched multi-armed bandits, batched local asymptotic normality.

*Assistant Professor, Department of Marketing, Virginia Tech. e-mail: yanx19@vt.edu.
†Assistant Professor, Department of Economics, Virginia Tech. e-mail: bzhou@vt.edu.

1 Introduction

Network formation plays a fundamental role in various real-world applications, where establishing effective connections is key to optimizing overall outcomes. Many such applications can be represented as *graph models*, where *nodes* represent agents, *edges* indicate connections between them, and *weights* quantify the resulting edge-level production outcomes. For example, a manager overseeing a group of workers forms teams to complete tasks, with team-level outputs captured as edge weights. Similarly, an online social platform may recommend connections between users, where the success or failure of these recommendations in establishing social ties is represented by binary weights. In both cases, the goal of the manager or platform is to strategically form edges—whether by assembling effective teams or facilitating meaningful user connections—to maximize overall production outcomes. This requires identifying agent pairs with strong complementarities, ensuring better task execution or higher success rates in social interactions.

To achieve this, the decision-makers—such as firms, governments and platforms— must first collect the edge-level production outcome data to infer the agents' latent types and their complementarities (the *exploration* phase). Once sufficient information has been collected, a network formation policy can be updated and implemented to improve overall outcomes (the *exploitation* phase). However, performing both phases in a single round—often referred to as an *offline* policy—can be problematic. Updating too early with insufficient data leads to inaccurate estimates, while updating too late misses critical opportunities for improvement. Both scenarios reflect an imbalance in the exploration-exploitation trade-off, resulting in inefficiencies and suboptimal outcomes.

A more effective approach is to use an algorithm that adaptively updates the policy as new outcome data become available, balancing exploration and exploitation to maximize accumulated outcomes. In this paper, we propose such an adaptive or *online* policy aimed at forming more productive networks. Specifically, we frame it as a *Batched Multi-Armed Bandit (Batched MAB)* problem, where each possible pair of agent latent types is treated as an arm to pull. Our algorithm runs in batches: Before each batch, we determine a network formation policy by solving a maximum-weight matching problem based on estimates from previously collected data. We then implement the policy for the batch to improve production and, simultaneously, collect the generated data. This newly collected data is used to further refine production estimates and update the policy for the next batch. By repeating this process iteratively, we enable continuous improvement in network formation decisions over time.

One challenge decision-makers often face is that only edge-level outcomes—such as teamwork

output or the connection between two social-network users—are observable, while individual contributions typically remain unobserved. In many contexts, such as teamwork (Weidmann and Deming (2021)), medical matching (Agarwal (2015)), or online dating (Lee (2016)), researchers have found that individual characteristics—such as a worker's soft skills or a user's personality—that are critical to edge-level outcomes cannot be directly observed. This limitation makes it difficult to assess the impact of each agent's unobserved heterogeneous characteristics and the complementarities or interplay between them. To address this, we build on the nonlinear network framework for team production developed by Bonhomme (2021), which is based on a weighted stochastic block model (WSBM). In this framework, agents' unobserved heterogeneous characteristics are captured by discrete latent types. The parameter determines the edge-level production outcome, denoted by $\boldsymbol{\theta}$, depends on the latent types of the corresponding agent pair. A key advantage of this WSBM-based framework is that it imposes no functional form restrictions on how agents' latent types influence outcomes, allowing for more flexible modeling of heterogeneous complementarities compared to production models with functional form assumptions, such as strongest/weakest-link specifications (see, e.g., Johari et al. (2018b)) or constant elasticity of substitution (see, e.g., Ahmadpoor and Jones (2019)).

Our estimation procedure, applied separately to each batch, is based on a *mean-field variational approximation* for the intractable likelihood of the WSBM model (see Bonhomme (2021) and Graham and De Paula (2020, Chapter 5)). To analyze its asymptotic properties, we build on Le Cam's asymptotic theory of limit experiments, particularly the recent extension by Hirano and Porter (2025) for adaptive experiments and batched bandits. Specifically, we prove a *batched local asymptotic normal (BLAN)* result for the WSBM-based production model under different pre-determined network-forming policies for each batch. This result states that the complex network production experiment within each batch can be approximated by a much simpler *limit experiment* involving a single Gaussian observation. Consequently, our T-batch network experiment can be approximated by a T-horizon Gaussian bandit environment in the sense of Hirano and Porter (2025). That is, we observe T Gaussian observations, each generated by a different normal distribution associated with a distinct policy. These observations serve as maximum likelihood estimates (MLEs) in the limit experiment, one from each batch. In addition, leveraging the techniques of Bickel et al. (2013), we show that the estimates based on the likelihood variational approximation asymptotically converge to these limiting MLEs, ensuring their consistency and asymptotic normality. Simultaneously, the variational estimation provides a *local variational approximation (LVA)* for each latent type vari-

able in the form of a discrete probability distribution, which quickly converges to the true latent type. All these results are developed under a *batched sparse network asymptotics*, which reflects real-world constraints by assuming that within each batch, the expected number of edges of each node (or *degree*) grows at most at the logarithmic rate of the number of nodes.

The theoretical results above convey a simple yet powerful message: the T batched variational estimates for the production parameter $\boldsymbol{\theta}$ can be treated as Gaussian signals. Building on this insight, we design our online network formation policy. Specifically, we apply Bayesian updating to obtain a posterior distribution for $\boldsymbol{\theta}$ using all completed batches. Similarly, we update the batched LVAs in the form of multinomial distributions to derive the posterior distributions of agents' latent types. We then compute a point estimate for $\boldsymbol{\theta}$ by taking the posterior mean and generating realizations of agents' latent types by sampling from the LVA posteriors. With these estimates and realizations, we then formulate and solve a (possibly constrained) matching problem to maximize the edge weights (or outcomes), which determines the network formation policy for the next batch. Simulation results show that our algorithm can substantially improve overall outcomes within just a few batches. Moreover, our adaptive experiment yields increasingly accurate estimates for both the production parameter $\boldsymbol{\theta}$ and the agents' latent types over time and achieves significant gains (i.e., lower regret). Finally, we show that the algorithm remains robust in non-stationary environments, where agents may enter and exit the network.

This paper builds on the growing literature on WSBM-based network models (Bonhomme (2021), Jochmans (2024)), which capture nonlinear complementarities—a key feature documented in recent empirical studies, such as Bergeron et al. (2022) in the context of tax collector teams and Xu et al. (2024) in the context of salesforce teams. Our study contributes to this field by establishing the asymptotic equivalence of variational estimation and the maximum likelihood estimation under sparse network asymptotics, drawing on techniques from Bickel and Chen (2009) for the original stochastic block models (Holland et al. (1983)). Additionally, we contribute to the multi-armed bandit literature, particularly on online policies for team formation and assignment problems (Johari et al. (2018a), Johari et al. (2021), Eichhorn et al. (2022)). Incorporating WSBM into this setting introduces two key departures from standard MAB formulations: (i) the arms—here, type pairs—are not directly observed when pulled, but only their probabilities (via LVAs), and (ii) instead of relying on descriptive statistics, the network formation policy is determined by solving an optimization problem.To address these challenges, we adopt the batched MAB framework (Zhang et al. (2020)) and leverage Le Cam's asymptotic theory for inference in batched settings (Hirano

and Porter (2025)), enabling the integration of complex network econometric modeling into online decision-making.

The rest of the paper is organized as follows. Section 2 sets up the network production model, formulates our network formation task as a Batched MAB problem, and presents the online policy. Section 3 details our estimation strategy, while Section 4 establishes the asymptotic results. Section 5 discusses the Bayesian updating procedure across batches in detail. Section 6 carries out a Monte Carlo study. Finally, Section 7 concludes the paper.

2 Online Network Formation as Batched Multi-armed Bandits

2.1 Setup

Suppose we have n agents, labeled $1, \ldots, n$, represented as *nodes* in a graph. Each agent i belongs to one of K types. The unobserved type of agent i is recorded in a (latent) random variable Z_i, which takes values in $[K] \equiv \{1, \ldots, K\}$. We assume the Z_i's are independently and identically distributed (i.i.d.) with probabilities given by

$$P(Z_i = a) = \pi(a), \tag{2.1}$$

for $a = 1, \ldots, K$, where $\pi(1) + \cdots + \pi(K) = 1$.

Agents are assigned into pairs — represented as *edges* in the graph — to complete m single-type tasks. These collaborations are defined by mappings $i_1(j)$ and $i_2(j)$ from $\mathcal{M} \equiv \{1, \ldots, m\}$ to $\mathcal{N} \equiv \{1, \ldots, n\}$, representing the two members in a pairwise connection j.[1] Throughout this paper, we use "group" and "connection" interchangeably.

The outcome of the Task j, completed by Group j, is modeled as the *weight* of the edge in the graph and is measured by a random variable $Y_j \in \mathcal{Y}$. The weight Y_j may be binary (e.g., when $Y_j = 1$ for success and $Y_j = 0$ for failure, with $\mathcal{Y} = \{0, 1\}$) or continuous ($\mathcal{Y} = \mathbb{R}$). We assume that Y_j's are i.i.d. with distribution

$$p_\theta \left(Y_j \,|\, Z_{i_1(j)}, Z_{i_2(j)} \right), \tag{2.2}$$

[1] In this paper, we focus on two-member connections, i.e., pairwise edge. Extending to M-member edges ($M > 2$) would require notions of hypergraphs, and the asymptotic results developed in Section 3 would need to be generalized accordingly. We leave this extension for future work.

which depends on the type pair $\left(Z_{i_1(j)}, Z_{i_2(j)}\right)$ of Group j, where $\theta = \theta_{Z_{i_1(j)} Z_{i_2(j)}} \in \mathbb{R}$ is an unknown parameter.[2] Here, we assume $i_1(j) \neq i_2(j)$ for all j, as we consider only two-member groups. We impose no functional form restrictions on the probability distribution p_θ, aside from some mild regularity conditions .

Our goal is to design an algorithm that optimally assigns agents into a network with pairwise connections to maximize overall outcomes.

Remark 1. *In this paper, we do not consider different types of tasks (i.e., task heterogeneity) or the matching problem between different types of agents and different types of tasks. For studies addressing this aspect, see, e.g., Johari et al. (2021).*

Remark 2. *Although we do not impose it in this paper, one could additionally assume the symmetry condition $p_\theta(\cdot \mid a, b) = p_\theta(\cdot \mid b, a)$ for all $a, b \in [K]$, implying that positions in a pair are interchangeable.*

2.2 Network Formation as an Optimization Problem

We aim to maximize the overall expected production

$$\max_{\psi \in \Psi} \psi(a, b)\, \Upsilon(a, b) \tag{2.3}$$

where $\Upsilon(a, b) \equiv \mathbb{E}\left(Y_j \mid Z_{i_1(j)} = a, Z_{i_2(j)} = b\right)$ represents the expected production of a type a-b pair.[3] The function $\psi(a, b)$ specifies the *pairing policy*, which determines the *frequency* or *probability mass* with which each type pair a-b is assigned across m pairs. The policy space Ψ to which ψ belongs, without further restrictions introduced, is defined as

$$\Psi = \left\{ A \in [0,1]^{K \times K} \;\middle|\; \sum_{a=1}^{K} \sum_{b=1}^{K} A[a, b] = 1 \right\}. \tag{2.4}$$

This ensures that each $\psi(a, b)$ is within $[0, 1]$ and that the summation of all probabilities equals one. In real-world applications, the optimization problem often involves additional constraints.

[2] For simplicity, we focus on the univariate case where θ_{ab} is scalar for each type pair a-b, $a, b \in [K]$. However, our results can be straightforwardly extended to the multivariate case, though at the cost of involving more complex notation.

[3] Alternative production evaluation measures (e.g., variance-adjusted metrics) can also be considered, with the optimization problem adjusted accordingly.

One important constraint is the workload constraint, discussed in works such as Bergeron et al. (2022) and Xu et al. (2024), which ensures fair task distribution among agents. Another frequently imposed constraint in the adaptive experiment literature is the so-called "clipping constraint," which prevents any arm-pulling probability from converging to 0, ensuring that each arm is explored with a non-negligible probability. These constraints are formally stated as follows.

- *The workload constraint:* In practice, each agent can only participate in a limited number of tasks within a specific period. Let $[d_l, d_h]$, with $d_l, d_h \in \mathbb{N}^+$ and $d_l < d_h$, denote the range of tasks any single agent can undertake. This agent-level workload constraint translates into the following policy-level workload constraint:

$$\sum_{b\in[K]} (\psi(a,b) + \psi(b,a))m \in [n_a d_l, n_a d_h], \quad \text{for all } a \in [K], \tag{2.5}$$

  where $n_a = \sum_{i\in\mathcal{N}} \mathbb{1}\{Z_i = a\}$ is the total number of type-a agents.

- *The clipping constraint:* We let $\psi(a,b)$ for each type combination a-b, $a, b \in [K]$, as arm selection probabilities, satisfy

$$\psi(a,b) \in [\lambda, 1-\lambda], \tag{2.6}$$

  for some clipping rate $\lambda \in (0, 0.5)$.

As a result, the optimization problem becomes

$$\max_{\psi\in\Psi^c} \psi(a,b)\, \Upsilon(a,b) \tag{2.7}$$

where $\Psi^c = \{\psi \in \Psi \,|\, \psi \text{ satisties (2.5) and (2.6)}\}$ is the constrained policy space. In an oracle scenario, where both agent types $\boldsymbol{Z} = (Z_1, \ldots, Z_n)'$ and the production distribution p_θ are fully known, this optimization problem can be solved directly using linear programming. We let ψ^* denote the associated oracle solution.

2.3 Online Network Formation as Non-standard Batched MAB Problem

However, in practice, neither agent types $\boldsymbol{Z}$ nor the production distribution p_θ are directly observable. To address this, one must first estimate them using sampled pairing and production data (exploration); then, based on these estimates, optimize pair assignments by pairing members who

complement each other most effectively (exploitation). This process, when performed in a single iteration, is often referred to as the *offline policy learning*. The manager or platform adopts a fixed policy afterwards and does not collect additional data to refine it further. Alternatively, a more effective strategy is to conduct adaptive experiments, where the agent-pairing policy is sequentially updated based on continuously collected data. This iterative process, known as *online policy learning*, strikes a balance between exploration (gathering data to improve estimates) and exploitation (leveraging current estimates to maximize performance outcomes).

To achieve this, we frame the iterative learn-and-earn process as a *Multi-Armed Bandits (MAB)* framework and design an algorithm tailored specifically for adaptive network formation. In the standard MAB problem, the exploration-exploitation tradeoff is modeled as a gambler choosing between multiple slot machine arms, each with an unknown payout distribution, in order to maximize total rewards over time. The gambler must balance experimenting with different arms to discover the most rewarding one and exploiting the arm that appears to offer the highest return. Building on this analogy, we treat all possible agent-type combinations as the arms of slot machines in our network formation problem. However, real-world network formation problem is inherently more complex and differs from the standard MAB framework in two major aspects.

First, in addition to the production distribution p_θ, the agent types $\boldsymbol{Z}$ are also unobserved. This makes the problem nonstandard because, when assigning pair members, the exact type combination is unknown — meaning the specific arm being pulled is unknown. Instead, we work with probabilities of $\boldsymbol{Z}$, estimated via variational approximation (see Section 3). These probabilities, in turn, determine the likelihood of each type combination being selected as an arm. As more data is gathered, the probability distributions of $\boldsymbol{Z}$ become increasingly concentrated around the latent true types, enhancing the manager's understanding of each agent's type.

Second, the sampling policy, which determines the arm-pulling probability for the next round, differs significantly from that of a standard bandit problem. In a typical bandit setting, the policy is often an explicit function of descriptive statistics, such as accumulated rewards and arm-pulling frequencies. In contrast, our network formation policy requires solving the optimization problem in (2.3), with unknown parameters replaced by the most recent estimates. Furthermore, as mentioned above, the optimization must also account for practical constraints, such as the workload constraint in (2.5), adding substantial complexity to the problem.

To tackle these complexities, we design our algorithm as a nonstandard *batched MAB* problem, where the network formation policy ψ is updated in discrete batches rather than continuously.

Specifically, we run the process for T batches. In each batch $t \in \{1, \ldots, T\}$, we pair n agents into m_t pairs (to complete m_t tasks) based on the policy $\psi_t(a, b)$. At the end of this batch, the outcomes of these tasks are recorded in the random variables $Y_{t,j}$ for $j = 1, \ldots, m_t$. Both T and m_t are predetermined or treated as exogenous in our framework, making them deterministic.We require ψ_t be specified based on the information available before each batch t. Formally, let $\boldsymbol{Y}_s = (Y_{s,1}, \ldots, Y_{s,m_s})'$, and let the accumulation of information before batch t be captured by

$$\mathcal{F}_t = \sigma\left(\boldsymbol{Y}_s, \psi_s, s = 1, \ldots, t-1\right).$$

We impose the following assumption.

Assumption 1 (Pre-determined Policy)**.** *For each batch $t = 1, \ldots, T$, we assume that the policy $\psi_t \in \Psi^c$ is adaptive to $\mathcal{F}_t$.*

The policy pre-specification condition in Assumption 1 naturally applies to all feasible network formation algorithms, as these can only utilize the data available at the time. We nevertheless explicitly state this as a formal assumption to emphasize this point and to indicate the policy ψ_t remains fixed throughout batch t. Moreover, this condition is also essential for the estimation procedure, as it ensures the *Network Exogeneity* condition outlined in Bonhomme (2021, Assumption 1). This condition prevents scenarios where agents have prior knowledge of pair-specific shocks before forming a pair.

Our online network formation algorithm is outlined as follows:

Algorithm 1 (Online Network Formation Policy)**.** *To start, pairs for batch $t = 1$ are assigned randomly, subject to the workload constraint. The following steps are then iteratively performed for $t = 1, \ldots, T-1$:*

1) *Estimation: Based on $\boldsymbol{Y}_t$, estimates for agent types and production distribution are obtained via procedures in Section 3, denoted as $\hat{\boldsymbol{\theta}}_t$ and $\hat{\boldsymbol{Z}}_t$, respectively.*

2) *Updating: Bayesian updates in Section 5 are applied to the estimates, treating each batch's estimates as individual signals. The updated estimates are denoted by $\hat{\boldsymbol{\theta}}_{1:t}$ and $\hat{\boldsymbol{Z}}_{1:t}$.*

3) *Optimization: The paring policy ψ_{t+1} is obtained by solving the optimization problem in (2.7), with $\boldsymbol{\theta}$ and $\boldsymbol{Z}$ replaced by $\hat{\boldsymbol{\theta}}_{1:t}$ and $\hat{\boldsymbol{Z}}_{1:t}$, respectively.*

We will provide a detailed discussion of the estimation step using variational approximation, along with the updating step, in the next section.

3 Estimation via Variational Approximation

In this section, we present our estimation strategy using a variational approximation technique and establish the corresponding asymptotic results. The estimation procedure is applied independently to each batch t based on the batch data $\boldsymbol{Y}_t$ for $t = 1, \ldots, T$, and so is the asymptotic analysis, which is conducted by letting $m_t \to \infty$ as $n \to \infty$ within batches. For clarity and simplicity, we omit the batch index t in the subsequent discussion when no confusion arises.

3.1 The Mean-field Approximation

Recall the individual production model, $p_\theta(\cdot|a, b)$, for the type pair a-b. Under the i.i.d. assumption for Y_j, the conditional distribution of $\boldsymbol{Y} = (Y_1, \ldots, Y_m)'$ given $\boldsymbol{Z}$ is

$$p_{\boldsymbol{\theta}}(\boldsymbol{Y}|\boldsymbol{Z}) = \prod_{j=1}^{m} p_\theta(Y_j|Z_{i_1(j)}, Z_{i_2(j)}), \tag{3.1}$$

where vector-valued $\boldsymbol{\theta} \in \mathbb{R}^{K^2}$ collects the individual pair production parameters $\theta = \theta_{ab}$ for $a, b \in [K]$. The likelihood of $(\boldsymbol{Y}, \boldsymbol{Z})$ is then given by

$$f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{Z}) = p_{\boldsymbol{\pi}}(\boldsymbol{Z}) p_{\boldsymbol{\theta}}(\boldsymbol{Y}|\boldsymbol{Z}), \tag{3.2}$$

where $p_{\boldsymbol{\pi}}(\boldsymbol{Z}) = \prod_{i=1}^{n} \pi(Z_i)$ represents the marginal likelihood of $\boldsymbol{Z}$. This model is referred to as the *complete model*, under the assumption that $\boldsymbol{Z}$ is observed.[4]

However, in practice, $\boldsymbol{Z}$ is unobserved. The marginal distribution of $\boldsymbol{Y}$ is obtained by summing over all configurations of $\boldsymbol{e} \in [K]^n$:

$$g_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}) = \sum_{\boldsymbol{e} \in [K]^n} f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{e}). \tag{3.3}$$

Computing this summation over $\boldsymbol{e} \in [K]^n$ requires K^n operations, which quickly becomes computationally infeasible as n grows large. To address this challenge, we employ a *mean-field variational approximation*, a widely adopted technique in the stochastic block model literature, as detailed below.

[4] This aligns with the concept of the *Complete Graph Model (CGM)* introduced in Bickel et al. (2013) for the classical stochastic block model.

We begin by expressing the joint density as $f(\boldsymbol{Y}, \boldsymbol{Z}) = h(\boldsymbol{Z}|\boldsymbol{Y})g(\boldsymbol{Y})$, where $h(\boldsymbol{Z}|\boldsymbol{Y})$ represents the conditional density of $\boldsymbol{Z}$ given $\boldsymbol{Y}$. The mean-field variation approach approximates the intractable conditional density $h(\boldsymbol{Z}|\boldsymbol{Y})$ with a factorized distribution $q(\boldsymbol{Z}) \equiv \prod_i q_i(Z_i)$. Here, each q_i represents an individual approximation of the type probabilities for the latent variable Z_i, referred to as a *local variational approximation*. The goal is to make $q(\boldsymbol{Z})$ as close as possible to $h(\boldsymbol{Z}|\boldsymbol{Y})$, of which the difference between the two distributions is measured using the Kullback–Leibler (KL) divergence:

$$D_{\mathrm{KL}}\left(q(\boldsymbol{Z}) \,\|\, h(\boldsymbol{Z}|\boldsymbol{Y})\right) \equiv \sum_{\boldsymbol{Z}\in[K]^n} q(\boldsymbol{Z}) \log \frac{q(\boldsymbol{Z})}{h(\boldsymbol{Z}|\boldsymbol{Y})}. \tag{3.4}$$

Expanding this expression yields:

$$\begin{aligned} D_{\mathrm{KL}}\left(q(\boldsymbol{Z}) \,\|\, h(\boldsymbol{Z}|\boldsymbol{Y})\right) &= \sum_{\boldsymbol{Z}\in[K]^n} q(\boldsymbol{Z}) \left[\log q(\boldsymbol{Z}) - \log h(\boldsymbol{Z}|\boldsymbol{Y})\right] \\ &= \sum_{\boldsymbol{Z}\in[K]^n} q(\boldsymbol{Z}) \left[\log q(\boldsymbol{Z}) - \log f(\boldsymbol{Y}, \boldsymbol{Z}) + \log g(\boldsymbol{Y})\right] \\ &= \log g(\boldsymbol{Y}) - J(\boldsymbol{Y}), \end{aligned} \tag{3.5}$$

with

$$J(\boldsymbol{Y}) = J_{\boldsymbol{\pi},\boldsymbol{\theta};\boldsymbol{q}}(\boldsymbol{Y}) \equiv \sum_{\boldsymbol{Z}\in[K]^n} q(\boldsymbol{Z}) \left[\log \frac{f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{Z})}{q(\boldsymbol{Z})}\right]$$

The matrix-valued $\boldsymbol{q} \in \mathbb{R}^{K\times n}$ collects the variational approximations for individual latent variables, where each column $\boldsymbol{q}_i \equiv (q_1(1), \ldots, q_i(K))'$, for $i = 1, \ldots, n$, is referred to as a *variational approximation (LVA)*.

The term $J(\boldsymbol{Y})$ is referred to as the *Evidence Lower Bound (ELBO)*. Since the KL-divergence is always non-negative, the ELBO satisfies $J(\boldsymbol{Y}) \leq \log g(\boldsymbol{Y})$, meaning it serves as a lower bound on the log-likelihood (with $\boldsymbol{Y}$ being the "evidence"). The ELBO is particularly advantageous due to its tractability. Specifically, it simplifies to

$$\begin{aligned} J_{\boldsymbol{\pi},\boldsymbol{\theta};\boldsymbol{q}}(\boldsymbol{Y}) = &\sum_{i=1}^{n}\sum_{a=1}^{K} q_i(a)\left(-\log q_i(a) + \log \pi(a)\right) \\ &+ \sum_{j=1}^{m}\sum_{a=1}^{K}\sum_{b=1}^{K} q_{i_1(j)}(a) q_{i_2(j)}(b) \log p_\theta(Y_j|a, b), \end{aligned} \tag{3.6}$$

where the summations over i and j require only $n \times K + m \times K^2$ operations in total.

To minimize the distance between $J(\boldsymbol{Y})$ and $\log g(\boldsymbol{Y})$, one can maximize $J_{\boldsymbol{\pi},\boldsymbol{\theta};\boldsymbol{q}}(\boldsymbol{Y})$ by optimizing the collection of type probabilities $\boldsymbol{q}$ for fixed values of $(\boldsymbol{\pi},\boldsymbol{\theta})$. This optimization is embedded in the *Expectation-Maximization (EM)* algorithm (see, e.g., Daudin et al. (2008)), which proceeds as follows:

Algorithm 2 (Expectation-Maximization)**.**

- *Initialization: Specify initial values for all parameters, denoted* $(\dot{\boldsymbol{\pi}},\dot{\boldsymbol{\theta}},\dot{\boldsymbol{q}})$.

- *Expectation Step (E-Step): Given* $(\dot{\boldsymbol{\pi}},\dot{\boldsymbol{\theta}})$*, update* $\dot{\boldsymbol{q}}$ *by solving*

$$\dot{\boldsymbol{q}} = \arg\max_{\boldsymbol{q}} J_{\dot{\boldsymbol{\pi}},\dot{\boldsymbol{\theta}};\boldsymbol{q}}(\boldsymbol{Y}); \tag{3.7}$$

- *Maximization Step (M-Step): Given* $\dot{\boldsymbol{q}}$*, update* $(\dot{\boldsymbol{\pi}},\dot{\boldsymbol{\theta}})$ *by solving*

$$(\dot{\boldsymbol{\pi}},\dot{\boldsymbol{\theta}}) = \arg\max_{\boldsymbol{\pi},\boldsymbol{\theta}} J_{\boldsymbol{\pi},\boldsymbol{\theta};\dot{\boldsymbol{q}}}(\boldsymbol{Y}); \tag{3.8}$$

- *Exit: Iteratively apply the E-Step and M-Step until convergence. Denote the final outputs by* $(\check{\boldsymbol{\pi}},\check{\boldsymbol{\theta}},\check{\boldsymbol{q}})$*, referred to as the variational estimates.*

4 Asymptotic Results

For our theoretical analysis, we introduce a *batched sparse network asymptotics* framework. This builds on the sparse network asymptotics commonly used to study the classical stochastic block model, which, unlike *dense network asymptotics*, provides accurate approximations for sparse networks (see, e.g., Bickel et al. (2013), Amini et al. (2013)). Specifically, it assumes that the expected number of connections per node, or *degree*, does not grow faster than $O(\log n)$. This assumption aligns with empirical observations in large networks datasets, where the number of links per node is typically small relative to the total number of nodes and can be approximated as logarithmic in scale.

We extend this idea to our batched setting by allowing the number of connections (or tasks) m_t of each batch t to grow at the rate $O(n \log n)$ as the number of nodes (or agents) n increases. This choice is also motivated by real-world applications, such as the house-selling context studied by **?**, where each agent can only participate in a limited number of tasks within a given batch period (e.g., one month).

We formalize this in the following assumption.

Assumption 2 (Batched Sparse Network)**.** *For each batch $t = 1, \ldots, T$, we let $m_t \to \infty$ as $n \to \infty$ such that $m_t/(n \log n) \to \lambda_t$ for some $\lambda_t \in (0, \infty)$.*

4.1 LAN Result When the Agent Characteristics are Observed

We begin by establishing a *batched local asymptotic normality (batched LAN)* result for the complete model, where the latent types $\boldsymbol{Z}$ are assumed to be observed. To achieve this, we impose two high-level assumptions on the individual production model $p_\theta(\cdot|a, b)$: a *conditional LAN* condition, under which the likelihood ratio admits a second-order Taylor series type expansion with asymptotically negligible higher-order terms, and the asymptotic normality of the maximum likelihood estimator (MLE). These assumptions are formally stated as follows.

Assumption 3. *(a) (Conditional LAN). Let $\vartheta \in \mathbb{R}$, and suppose Y_j is generated by the production model $p_\theta(\cdot|a, b)$. For all $a, b \in [K]$, we assume the following expansion holds:*

$$\log \prod_{j=1}^{m} \frac{p_{\theta+\vartheta/\sqrt{m}}(Y_j|a, b)}{p_\theta(Y_j|a, b)} = \frac{\vartheta}{\sqrt{m}} \sum_{j=1}^{m} \dot{\ell}_\theta(Y_j|a, b) + \frac{\vartheta^2}{2} \mathcal{I}_\theta(a, b) + o_{\mathrm{P}}(1),$$

where $\dot{\ell}_\theta(\cdot|a, b)$ is the score function, satisfying $\mathbb{E}_\theta\big[\dot{\ell}_\theta(Y_j|a, b)\big] = 0$, and the Fisher information $\mathcal{I}_\theta(a, b) \equiv \mathbb{E}_\theta\big[\dot{\ell}^2_\theta(Y_j|a, b)\big]$ exists.

(b) (Asymptotically Normal MLE). For all $a, b \in [K]$, let θ^*_{ab} denote the maximum likelihood estimator (MLE). We assume that

$$\sqrt{m}(\theta^*_{ab} - \theta_{ab}) \sim \mathcal{N}(0, \mathcal{I}^{-1}_\theta(a, b)).$$

Assumption 3 is mild and has been shown to hold for many specific distributions. For continuous Y_j, this conditional LAN result follows directly from the *differentiable in quadratic mean (DQM)* condition. This condition states that there exist a measurable function $\dot{\ell}_\theta(Y_j|a, b)$ such that, as $\epsilon \to 0$,

$$\int_{-\infty}^{+\infty} \left(\sqrt{\frac{p_{\theta+\epsilon}(y|a, b)}{p_\theta(y|a, b)}} - 1 - \frac{\epsilon}{2} \dot{\ell}_\theta(y|a, b) \right)^2 p_\theta(y|a, b) \mathrm{d}y = o(\epsilon^2),$$

for all $a, b \in [K]$. The DQM condition holds if the density function $p_\theta(\cdot|a,b)$ is absolutely continuous with a first-order derivative $\dot{\ell}_\theta(\cdot|a,b)$. In this case, the score function is given by $\dot{\ell}_\theta(Y|a,b) = -\dot{p}_\theta(Y|a,b)/p_\theta(Y|a,b)$ and the Fisher information is $\mathcal{I}_\theta(a,b) = \mathbb{E}_\theta\big[\dot{\ell}^2_\theta(Y_j|a,b)\big]$. The asymptotic normality of the corresponding MLE is generally satisfied in smooth parametric models under mild regularity conditions; see, e.g., Van der Vaart (2000, Chapter 5). For binary Y_j, the LAN condition arises from a standard second-order Taylor expansion, while the asymptotic normality of the MLE follows directly from the central limit theorem for binomial experiments. Detailed derivations of these results for the individual production model can be found in the proofs of Lemmas 1–2 in Bickel et al. (2013), where the score function is given by $\dot{\ell}_\theta(Y_j|a,b) = Y_j - \theta_{ab}$, and the Fisher information is $\mathcal{I}_\theta(a,b) = \theta_{ab}(1-\theta_{ab})$.

Define $\omega(a) = \log \frac{\pi(a)}{1-\sum_{l=1}^{K-1}\pi(l)}$ for $a = 1, \ldots, K-1$, and let $\boldsymbol{\omega}$ denote the collection of all $\omega(a)$. Since this transformation is one-to-one, we will use the notations $f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z})$, $g_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y})$, and $J_{\boldsymbol{\pi},\boldsymbol{\theta};\boldsymbol{q}}(\boldsymbol{Y})$ interchangeably with $f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z})$, $g_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y})$, and $J_{\boldsymbol{\omega},\boldsymbol{\theta};\boldsymbol{q}}(\boldsymbol{Y})$, respectively. In the following proposition, we establish the batched LAN result for the complete model where $\boldsymbol{Z}$ is considered observed.

Proposition 1. *Let Assumption 1–3 hold.*

i) The complete log-likelihood ratio, $\Lambda_f^{(n,m)}(\boldsymbol{\nu},\boldsymbol{\vartheta}) \equiv \log f_{\boldsymbol{\omega}+\frac{\boldsymbol{\nu}}{\sqrt{n}},\boldsymbol{\theta}+\frac{\boldsymbol{\vartheta}}{\sqrt{m}}}(\boldsymbol{Y},\boldsymbol{Z}) - \log f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z})$, *can be decomposed as*

$$\begin{aligned}\Lambda_f^{(n,m)}(\boldsymbol{\nu},\boldsymbol{\vartheta}) = &\sum_{a=1}^{K-1}\nu(a)\Delta_{\boldsymbol{\omega}}^{(n)}(a) - \frac{1}{2}\sum_{a=1}^{K-1}\sum_{b=1}^{K-1}\nu(a)\mathcal{Q}_{\boldsymbol{\omega}}(a,b)\nu(b)\\ &+\sum_{a=1}^{K}\sum_{b=1}^{K}\left(\vartheta_{ab}\Delta_{\boldsymbol{\theta}}^{(m)}(a,b) - \frac{1}{2}\vartheta_{ab}^2\mathcal{Q}_{\boldsymbol{\theta}}(a,b)\right) + o_{\mathrm{P}}(1),\end{aligned} \tag{4.1}$$

where

$$\begin{aligned}\Delta_{\boldsymbol{\omega}}^{(n)}(a) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\mathbb{1}\{Z_i = a\} - \pi(a)\right),\\ \mathcal{Q}_{\boldsymbol{\omega}}(a,b) &= \begin{cases}\pi(a)(1-\pi(a)), & \text{if } a = b\\ \pi(a)\pi(b), & \text{if } a \neq b\end{cases}\\ \Delta_{\boldsymbol{\theta}}^{(m)}(a,b) &= \frac{1}{\sqrt{m}}\sum_{j=1}^{m}\mathbb{1}\{Z_{i_1(j)} = a, Z_{i_2(j)} = b\}\dot{\ell}_\theta(Y_j|a,b),\\ \mathcal{Q}_{\boldsymbol{\theta}}(a,b) &= \psi(a,b)\mathcal{I}_\theta(a,b).\end{aligned} \tag{4.2}$$

ii) Under $f_{\boldsymbol{\omega}+\frac{\boldsymbol{\nu}}{\sqrt{n}},\boldsymbol{\theta}+\frac{\boldsymbol{\vartheta}}{\sqrt{m}}}$*, we have*

$$\begin{pmatrix} \boldsymbol{\Delta}_{\boldsymbol{\omega}}^{(n)} \\ \boldsymbol{\Delta}_{\boldsymbol{\theta}}^{(m)} \end{pmatrix} \Rightarrow \begin{pmatrix} \boldsymbol{\Delta}_{\boldsymbol{\omega}} \\ \boldsymbol{\Delta}_{\boldsymbol{\theta}} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \boldsymbol{\nu} \\ \boldsymbol{\vartheta} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\mathcal{Q}}_{\boldsymbol{\omega}} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\mathcal{Q}}_{\boldsymbol{\theta}} \end{pmatrix} \right), \tag{4.3}$$

where $\boldsymbol{\Delta}_{\boldsymbol{\omega}}^{(n)}$*,* $\boldsymbol{\Delta}_{\boldsymbol{\theta}}^{(m)}$*,* $\boldsymbol{\mathcal{Q}}_{\boldsymbol{\omega}}$*, and* $\boldsymbol{\mathcal{Q}}_{\boldsymbol{\theta}}$ *are matrices whose elements are defined in (4.2). Consequently,*

$$\begin{aligned} \Lambda_f^{(n,m)}(\boldsymbol{\nu},\boldsymbol{\vartheta}) \Rightarrow \Lambda(\boldsymbol{\nu},\boldsymbol{\vartheta}) = & \sum_{a=1}^{K-1} \nu(a)\Delta_{\boldsymbol{\omega}}(a) - \frac{1}{2}\sum_{a=1}^{K-1}\sum_{b=1}^{K-1} \nu(a)\mathcal{Q}_{\boldsymbol{\omega}}(a,b)\nu(b) \\ & + \sum_{a=1}^{K}\sum_{b=1}^{K}\left(\vartheta_{ab}\Delta_{\boldsymbol{\theta}}(a,b) - \frac{1}{2}\vartheta_{ab}^2\mathcal{Q}_{\boldsymbol{\theta}}(a,b)\right). \end{aligned}$$

iii) Let $(\hat{\boldsymbol{\omega}}_f, \hat{\boldsymbol{\theta}}_f)$ *denote the MLEs based on* $f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z})$*, we have*

$$\begin{aligned} \sqrt{n}(\hat{\boldsymbol{\omega}}_f - \boldsymbol{\omega}) &\sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\mathcal{Q}}_{\boldsymbol{\omega}}^{-1}\right), \\ \sqrt{m}(\hat{\boldsymbol{\theta}}_f - \boldsymbol{\theta}) &\sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\mathcal{Q}}_{\boldsymbol{\theta}}^{-1}\right). \end{aligned}$$

This result implicitly follows from Assumption 1, which states that the network formation policy ψ_t is pre-specified and thus exogenous for each batch t. Consequently, the joint likelihood of $\boldsymbol{Y}$ factorizes into the product of individual edge-level likelihoods Y_j, due to the i.i.d. assumption. Proposition 1 then follows almost directly from the high-level conditions in Assumption 3 — namely, the LAN expansion and the asymptotic normality of the MLE for the individual production model $p_\theta(\cdot \mid a, b)$ — provided that the sparse network asymptotics in Assumption 2 hold. For completeness, we provide a more detailed proof in Appendix A.

The LAN result within each batch invokes the *Asymptotic Representation Theorem* (see Le Cam et al. (1972), Hájek (1970), Van der Vaart (2000)). This theorem states that the inferential problem under a sequence of complex experiments — in this case, the network production model estimation — can be approximated by a much simpler experiment. This simpler experiment, referred to as the *limit experiment*, has a likelihood ratio that coincides with the limiting likelihood ratio of the sequence of experiments (here, $\Lambda(\boldsymbol{\nu},\boldsymbol{\vartheta})$). In the context of this paper, the limit experiment corresponds to observing single realizations of random variables $\boldsymbol{\mathcal{Q}}_{\boldsymbol{\omega}}^{-1}\boldsymbol{\Delta}_{\boldsymbol{\omega}}$ and $\boldsymbol{\mathcal{Q}}_{\boldsymbol{\theta}}^{-1}\boldsymbol{\Delta}_{\boldsymbol{\theta}}$, which follow normal distributions $\mathcal{N}(\boldsymbol{\nu}, \boldsymbol{\mathcal{Q}}_{\boldsymbol{\omega}}^{-1})$ and $\mathcal{N}(\boldsymbol{\vartheta}, \boldsymbol{\mathcal{Q}}_{\boldsymbol{\theta}}^{-1})$, respectively. This experiment is often referred to

as the Gaussian shift experiment, where a straightforward derivation confirms that the associated likelihood ratio equals $\Lambda(\boldsymbol{\nu}, \boldsymbol{\vartheta})$. The Asymptotic Representation Theorem further states that for every sequence of statistics in original experiments that converges, the limit corresponds to a statistic in the limit experiment. This property enables us to first solve the inferential problem in the simpler limit experiment and then identify the corresponding finite-sample representation in the original experiments.

It is important to emphasize that these Gaussian shift experiments for different batches are not identical. Specifically, restoring the notation t for batch t, the normal distributions $\mathcal{N}(\boldsymbol{\vartheta}, \boldsymbol{\mathcal{Q}}_{\boldsymbol{\theta},t}^{-1})$ have different variances, which depend on the corresponding network formation or sampling policy ψ_t. Indeed, as observed in (4.2), the inverse variance is given by $\mathcal{Q}_{\boldsymbol{\theta},t}(a,b) = \psi_t(a,b)\mathcal{I}_\theta(a,b)$. This is intuitive—when $\psi_t(a,b)$ is larger, meaning more a-b type pairs are connected, the estimate for the corresponding production parameter θ_{ab} will be more precise (with smaller variance).

Across batches, given that the network formation policy is pre-specified for each batch (Assumption 1) and variational estimation is applied separately to each batched dataset, it follows that the batched likelihood ratios jointly converge to their limits: $\big(\Lambda_{f,1}^{(n,m)}(\boldsymbol{\nu}, \boldsymbol{\vartheta}), \ldots, \Lambda_{f,T}^{(n,m)}(\boldsymbol{\nu}, \boldsymbol{\vartheta})\big) \Rightarrow \big(\Lambda_1(\boldsymbol{\nu}, \boldsymbol{\vartheta}), \ldots, \Lambda_T(\boldsymbol{\nu}, \boldsymbol{\vartheta})\big)$. This invokes Theorem 2 of Hirano and Porter (2025), the multi-batched Asymptotic Representation Theorem, which states that any statistical decision rules based on converging statistics in the original batched network formation problem can always be represented by some corresponding decision rules in the limiting T-observation Gaussian shift experiments. Building on this insight, we develop our adaptive network formation rule in Section 5. Before that, we first establish that the variational estimates obtained via Algorithm 2 are asymptotically equivalent to the (intractable) maximum likelihood estimates.

4.2 LAN Result When the Agent Characteristics are Unobserved

We now consider the real-world setting where the manager cannot directly observe the latent agent types, $\boldsymbol{Z}$, which must also be learned. The following proposition establishes that the marginal model $g_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y})$ is asymptotically equivalent to that of the complete model $f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{Z})$ with a probability converging to one as the sample size grows. Consequently, the former maintains the same locally asymptotically normal expansion as the latter.

To establish this result, we impose additional regularity conditions on the production model, as specified in Assumptions 4 and 5. These conditions, along with the corresponding proof, are detailed in Appendix A. They are mild and can be readily verified once the distribution is specified.

For example, in the case of continuous Y_j, they follow naturally when f is assumed to be normal. In the case of binary outcomes Y_j, these conditions have been validated in Bickel et al. (2013) as part of their proof of Theorem 1.

Theorem 1. *Assume that $f_{\boldsymbol{\pi},\boldsymbol{\theta}}$ is unimodal in $(\boldsymbol{\pi},\boldsymbol{\theta})$, that $\boldsymbol{\theta}$ has no identical columns, and that Assumptions 1–5 are satisfied.*

(i) Let $\mathcal{S}(\boldsymbol{\pi},\boldsymbol{\theta})$ denote the equivalent set of parameters up to a relabeling of the latent classes $\{1,\ldots,K\}$. Then we have

$$g_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}) = \max_{(\boldsymbol{\pi}',\boldsymbol{\theta}')\in\mathcal{S}(\boldsymbol{\pi},\boldsymbol{\theta})} f_{\boldsymbol{\pi}',\boldsymbol{\theta}'}(\boldsymbol{Y},\boldsymbol{Z}) + o_P(1), \tag{4.4}$$

where the $o_P(1)$ term is uniform over $\boldsymbol{\pi}\in\mathbb{R}^K$ and $\boldsymbol{\theta}\in\mathbb{R}^{K^2}$.

(ii) Let $(\hat{\boldsymbol{\omega}}_g,\hat{\boldsymbol{\theta}}_g)$ denote the maximum likelihood estimates based on $g_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y})$, then $\mathcal{S}(\hat{\boldsymbol{\omega}}_g,\hat{\boldsymbol{\theta}}_g)$ contains an element $(\hat{\boldsymbol{\omega}}'_g,\hat{\boldsymbol{\theta}}'_g)$ satisfying

$$\sqrt{n}(\hat{\boldsymbol{\omega}}'_g-\boldsymbol{\omega}) \sim \mathcal{N}\left(\mathbf{0},\boldsymbol{\mathcal{Q}}_{\boldsymbol{\omega}}^{-1}\right),$$
$$\sqrt{m}(\hat{\boldsymbol{\theta}}'_g-\boldsymbol{\theta}) \sim \mathcal{N}\left(\mathbf{0},\boldsymbol{\mathcal{Q}}_{\boldsymbol{\theta}}^{-1}\right).$$

Define the log likelihood ratio of the network production model as $\Lambda_g^{(n,m)}(\boldsymbol{\nu},\boldsymbol{\vartheta}) \equiv \log g_{\boldsymbol{\omega}+\frac{\boldsymbol{\nu}}{\sqrt{n}},\boldsymbol{\theta}+\frac{\boldsymbol{\vartheta}}{\sqrt{m}}}(\boldsymbol{Y}) - \log g_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y})$. Under the batched sparse network asymptotic, we have

$$\Lambda_g^{(n,m)}(\boldsymbol{\nu},\boldsymbol{\vartheta}) \Rightarrow \Lambda(\boldsymbol{\nu},\boldsymbol{\vartheta}).$$

Finally, we turn to the mean-field variational approximation of the log likelihood $\log g(\boldsymbol{Y})$, $J(\boldsymbol{Y})$, as introduced in Section 3.1, which we employ for estimation. The following theorem establishes that the log-likelihood ratio based on $J(\boldsymbol{Y})$ follows the same limiting LAN expansion as the complete model.

Theorem 2. *Assume the same conditions as in Theorem 1 hold.*

(i) Define $\tilde{J}_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y}) = \max_{\boldsymbol{q}} J_{\boldsymbol{\omega},\boldsymbol{\theta};\boldsymbol{q}}(\boldsymbol{Y})$, we have

$$\tilde{J}_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y}) = \max_{(\boldsymbol{\omega}',\boldsymbol{\theta}')\in\mathcal{S}(\boldsymbol{\pi},\boldsymbol{\theta})} \log f_{\boldsymbol{\omega}',\boldsymbol{\theta}'}(\boldsymbol{Y},\boldsymbol{Z}) + o_P(1),$$

where the $o_P(1)$ term is uniform over $\boldsymbol{\pi}\in\mathbb{R}^K$ and $\boldsymbol{\theta}\in\mathbb{R}^{K^2}$.

(ii) Let $(\breve{\boldsymbol{\omega}}, \breve{\boldsymbol{\theta}})$ denote the mean-field variational estimates based on $\tilde{J}_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y})$, then $\mathcal{S}(\breve{\boldsymbol{\omega}}, \breve{\boldsymbol{\theta}})$ contains an element $(\breve{\boldsymbol{\omega}}', \breve{\boldsymbol{\theta}}')$ satisfying

$$\sqrt{n}(\breve{\boldsymbol{\omega}}' - \boldsymbol{\omega}) \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\mathcal{Q}}_{\boldsymbol{\omega}}^{-1}\right),$$
$$\sqrt{m}(\breve{\boldsymbol{\theta}}' - \boldsymbol{\theta}) \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\mathcal{Q}}_{\boldsymbol{\theta}}^{-1}\right).$$

5 Online Network Formation Algorithm

The theoretical findings in Section 4 convey a simple yet powerful message: the T-batched complex network production problem can be approximated by a series of T Gaussian shift experiments. In each experiment, indexed by $t \in \{1, \ldots, T\}$, a single signal is observed, following a Gaussian distribution with a mean governed by the production parameter $\boldsymbol{\theta}$ and a variance associated with the network formation policy ψ_t. Consequently, these signals correspond to maximum likelihood estimates. The variational estimates $\breve{\boldsymbol{\theta}}_t$ in the network production model asymptotically converge to these limiting MLEs, making them finite-sample representations of the Gaussian signals. Additionally, the local variational approximations (LVAs), $\boldsymbol{q}_i$, obtained from the variational estimation—one from each batch—serve as proxies for posterior distributions of agents' latent types. Note that, although the signals for $\boldsymbol{\theta}$ and latent types are independent across batches, they are not identically distributed due to the continuously updated sampling policy ψ_t. Therefore, rather than simply taking their averages, we employ Bayesian updating to aggregate information and refine the estimates of the unknown parameters using all completed batches.

The Bayesian updating on $\boldsymbol{\theta}$ is pairwise: For each a-b type pair, we denote the corresponding variational estimate as $\breve{\theta}_{t,ab}$ and its standard error by $\breve{se}_{t,ab}$ from batch t, where $a, b \in [K]$ and $t = 1, \ldots, T$. Let $\tilde{\mu}_{t,ab}$ and $\tilde{\sigma}_{t,ab}$ represent the posterior mean and standard deviation of θ_{ab} after batch t, respectively, with the prior distribution for $\theta_{t,ab}$ specified as $\mathcal{N}(\tilde{\mu}_{0,ab}, \tilde{\sigma}_{0,ab}^2)$. The Gaussian Bayesian updating is then via

$$\tilde{\mu}_t = \left(\frac{\tilde{\mu}_{t-1}}{\tilde{\sigma}_{t-1}^2} + \frac{\breve{\theta}_t}{\breve{se}_t^2}\right) \Big/ \left(\frac{1}{\tilde{\sigma}_{t-1}^2} + \frac{1}{\breve{se}_t^2}\right) \quad \text{and} \quad \tilde{\sigma}_t = \frac{\tilde{\sigma}_{t-1}^2 \breve{se}_t^2}{\tilde{\sigma}_{t-1}^2 + \breve{se}_t^2}, \tag{5.1}$$

for $t = 1, \ldots, T$, which applies to all a-b pair and thus we omit them in the subscripts for simplicity.

Turning to agents' latent types, recall the local variational approximation of batch t for agent i, denoted by $\boldsymbol{q}_{t,i} \in \mathbb{R}^K$, which is expressed in the form of a categorical distribution as the posterior

of the latent Z_i. Let $\tilde{\boldsymbol{q}}_{t,i}$ denote the posterior distribution of Z_i after batch t, with the prior distribution for Z_i specified as $\tilde{\boldsymbol{q}}_{0,i}$. The multinomial Bayesian updating procedure is then given by

$$\tilde{\boldsymbol{q}}_{t,i}[k] = \frac{\tilde{\boldsymbol{q}}_{t-1,i}[k]\check{\boldsymbol{q}}_{t,i}[k]}{\sum_{k\in[K]} \tilde{\boldsymbol{q}}_{t-1,i}[k]\check{\boldsymbol{q}}_{t,i}[k]}, \tag{5.2}$$

for $t = 1, \ldots, T$, $i = 1, \ldots, n$, and $k = 1, \ldots, K$.

Given these posterior distributions after batch t, specifically $\mathcal{N}(\tilde{\mu}_{t,ab}, \tilde{\sigma}^2_{t,ab})$ for θ_{ab} and $\tilde{\boldsymbol{q}}_i$ for Z_i, we have two potential approaches for determining the unknown parameters. The first option is a *greedy* strategy, where we obtain a point estimate by taking the expectation with respect to the associated posterior distribution. Alternatively, following the principle of *Thompson sampling* (originally introduced in Thompson (1933)), we can draw a random realization from the posterior distribution. In both approaches, these point estimates or realizations are then treated as the true parameter values, allowing us to maximize the expected reward for the next batch. Based on these strategies, we propose our *Hybrid Greedy-Thompson (HGT) algorithm* (greedy-in-$\boldsymbol{\theta}$, Thompson-in-$\boldsymbol{q}$) as follows.

Algorithm 3 (Hybrid Greedy-Thompson)**.** *For all $a, b = 1, \ldots, K$ and $i = 1, \ldots, n$, specify the priors as $\mathcal{N}(\tilde{\mu}_{0,ab}, \tilde{\sigma}^2_{0,ab}) = \mathcal{N}(0, \infty)$ and $\tilde{\boldsymbol{q}}_i = Uniform(1, 2, \ldots, K)$. Then, iteratively performed the following steps for $t = 1, \ldots, T$:*

1) Optimization: In the weight-maximum matching problem (2.7), replace θ_{ab} with $\tilde{\mu}_{t-1,ab}$ and Z_i with a random draw from $\tilde{\boldsymbol{q}}_{t-1,i}$. Solve this optimization to obtain the network formation policy ψ_t. Apply ψ_t for batch t and record the outcome data $\boldsymbol{Y}_t$.

2) Estimation: Based on $\boldsymbol{Y}_t$, obtain estimates $\check{\boldsymbol{\theta}}_t$ and $\check{\boldsymbol{q}}_t$ using Algorithm 2.

2) Updating: Perform Bayesian updates in (5.1) and (5.2) to obtain $\tilde{\mu}_{t,ab}$ and $\tilde{\boldsymbol{q}}_{t,i}$, respectively.

Although Bayesian updating is a natural choice in this non-i.i.d. setting, its fundamental advantage should not be overlooked: the ability to incorporate any available prior knowledge of the unknown parameters. This feature is particularly valuable in practice. For instance, when forming teams for tasks, a manager may already have some knowledge of her workers' abilities, such as from their application materials. Similarly, in social platform recommendations, users often provide information during registration that reveals aspects of their latent types. However, if no prior information is available, we can simply specify uninformative priors—an infinite-variance normal

distribution for θ_{ab} and discrete uniform distribution for Z_i—in which case the posterior reduces to the likelihood.

6 Simulation

In this section, we conduct a Monte Carlo study to evaluate our network formation policy. We simulate an environment where a manager assigns workers to teams to complete tasks. Specifically, we consider a pool of $n = 40$ workers, whose characteristics are unobserved by the manager. These characteristics are represented by $K = 3$ latent types, drawn from a multinomial distribution with probabilities $(0.4, 0.3, 0.3)$ for types 1, 2, and 3, respectively. Once assigned, these types remain fixed and are treated as unknown parameters.

Each team's success in completing a task is represented by a binary outcome. The probability of success for a team composed of an a-b type pair is denoted by $\boldsymbol{\theta}(a, b) \in [0, 1]$, with all such probabilities collected in the matrix $\boldsymbol{\theta} \in [0, 1]^{3\times 3}$. For simplicity, We specify a symmetric outcome parameter matrix $\boldsymbol{\theta}$ as

$$\boldsymbol{\theta} = \begin{pmatrix} 0.10 & 0.20 & 0.30 \\ 0.20 & 0.40 & 0.60 \\ 0.30 & 0.60 & 0.80 \end{pmatrix},$$

where the rows and columns of $\boldsymbol{\theta}$ are ordered by type 1, 2, and 3 from top to bottom and left to right, respectively. The symmetry of $\boldsymbol{\theta}$ implies that the roles of the two workers in a team are interchangeable.

We run our experiments over $T = 18$ batches, where in each batch, the 40 workers are assigned to 500 teams to complete 500 tasks, averaging 25 tasks per worker. To better reflect real-world conditions, we impose a *fairness constraint* in the optimization problem (2.7), ensuring that each worker is assigned at least 20 times and at most 30 times per batch.

Figure 1 displays the histograms of the posterior means for each θ_{ab}, i.e., the $\tilde{\mu}_{t,ab}$'s, for $a, b = 1, 2, 3$, at Weeks 1, 8, and 18. As expected, these histograms become narrower over time, indicating that our algorithm effectively refines the parameter estimates through Bayesian updating. However, since the likelihood of the WSBM is often susceptible to multiple local optima (see, e.g., Nowicki and Snijders (2001), Bickel et al. (2013), Bonhomme (2021)), we sometimes encounter biased estimates, even though we initialize the variational EM algorithm from different starting points.

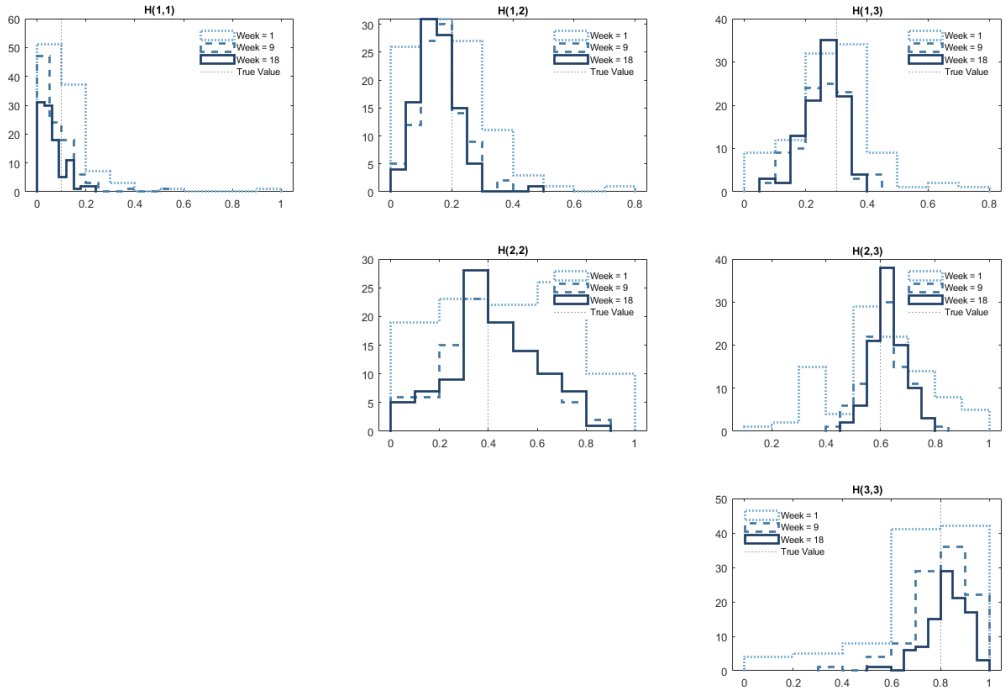


Figure 1: Histograms of posterior means for $\boldsymbol{\theta}$ based on batched variational estimates after Week 1 (light blue), Week 8 (blue), and Week 18 (dark blue), based on 100 repetitions.

Figure 2 presents the false labeling rates, defined as the proportion of incorrectly estimated Z_i values across 40 workers: $\sum_{i=1}^{40} \mathbb{1}\{\hat{Z}_i \neq Z_i\}/40$, where $\hat{Z}_i$ denotes the estimated latent type for worker i. We estimate Z_i as $\hat{Z}_i = \arg\max_k \tilde{\boldsymbol{q}}_i(k)$ for $k = 1, 2, 3$, recalling that $\tilde{\boldsymbol{q}}_i$ is the posterior distribution after batch t. The figure shows a significant decrease in the false labeling rate over a few batches, dropping from nearly 40% to about 20%, and further reducing to around 10% after 18 batches.

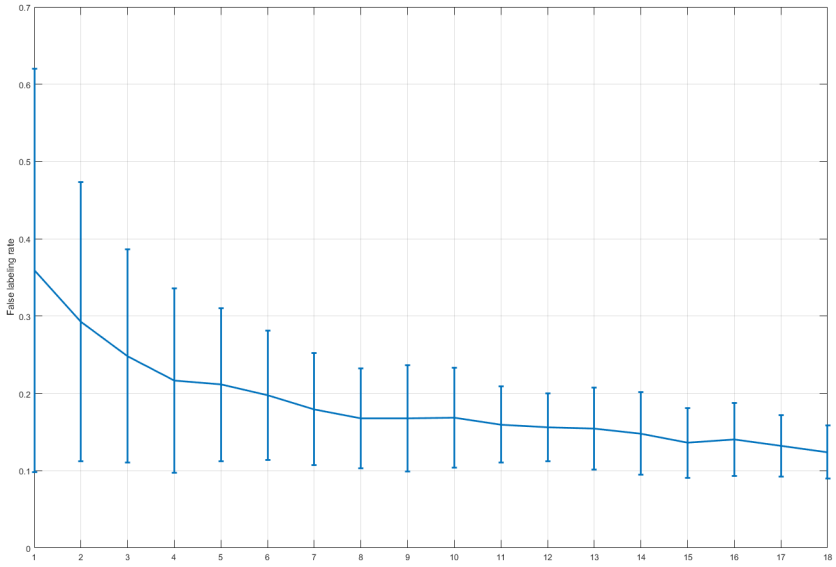


Figure 2: False labeling rates of 40 workers, defined as $\sum_{i=1}^{40} \mathbb{1}\{\hat{Z}_i \neq Z_i\}/40$, where the latent type estimate for worker i is given by $\hat{Z}_i = \arg\max_k \tilde{\boldsymbol{q}}_i(k)$ over $k = 1, 2, 3$.

Figure 3 illustrates the performance of our algorithm in a stationary environment, where no new agents enter the network, and no existing agents exit. We observe that the overall output improves significantly within several batches, approaching the oracle scenario in which the network production output is fully optimized as if the production parameter $\boldsymbol{\theta}$ were known and agents' latent types $\boldsymbol{Z}$ were observed.

Next, we demonstrate that our algorithm is also effective in a non-stationary environment, where new agents may enter, and existing agents may exit over time. Adaptive algorithms are particularly advantageous in such settings, as the network formation policy continuously updates with new data. In contrast, an offline policy, which relies on a single round of data collection, can be problematic since it remains fixed and may not be optimal for newly joined agents. To simulate this scenario, we consider a non-stationary case in which, among the 40 positions, some workers may leave and be replaced by new hires. As a result, the latent type associated with a given position may change with a certain probability and must be re-estimated.

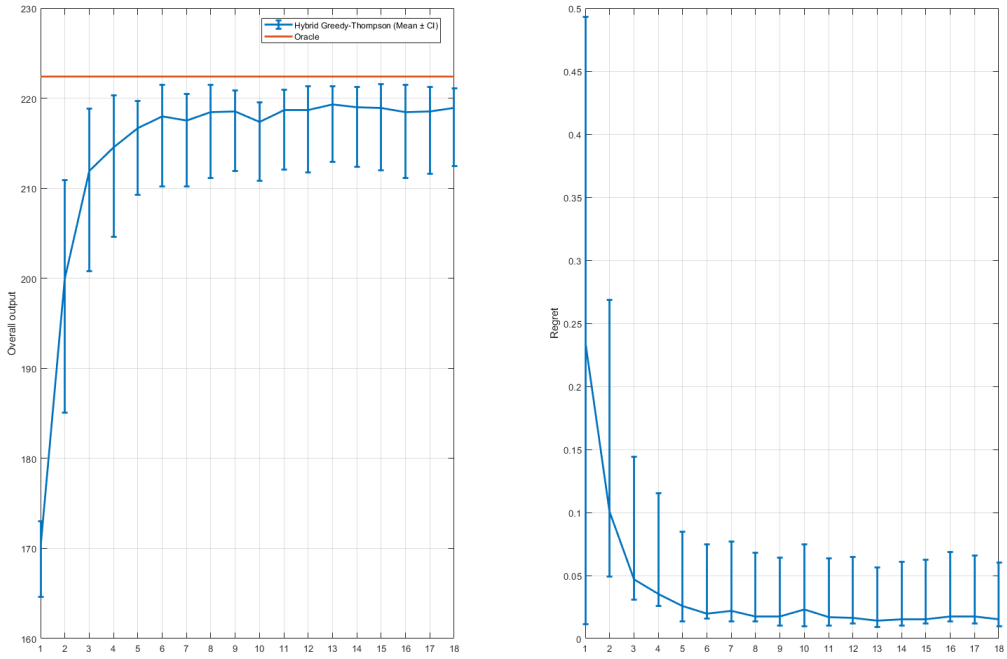


Figure 3: Average network production performance of our Hybrid Greedy-Thompson (HGT) algorithm (blue) compared to the oracle benchmark (orange), measured by the total number of successful tasks (left) and regret (percentage, right). Standard deviations (blue bars) are calculated over 100 repetitions.

Figure 4 presents the regret curves for four scenarios. In each scenario, a certain percentage of agents leave the network each period, and an equal number of new agents enter. This setup mimics real-world dynamics, such as job turnover in firms or user turnover in matching markets. We vary the turnover rate from low to high (i.e., from 0 to 10 percent) and plot the corresponding regret curves. The results show that even in a non-stationary environment, our algorithm effectively minimizes regret, adapting to changing network conditions over time.

This setting is particularly relevant in industries with high turnover rates, such as retail, customer service, and gig economy platforms, where employee or user churn is frequent. Traditional offline policies, which rely on a fixed exploration phase followed by an exploitation phase, face significant challenges in such environments. Specifically, by the time the policy reaches the exploitation phase, many of the initially explored agents may have already left, while newly arrived agents remain unexplored, leading to suboptimal performance. In contrast, our algorithm continuously adapts to these real-world dynamics by integrating exploration and exploitation in an ongoing

manner. This adaptability enables it to effectively design and sustain productive networks despite workforce or user base fluctuations, making it well-suited for practical applications where stability cannot be assumed.

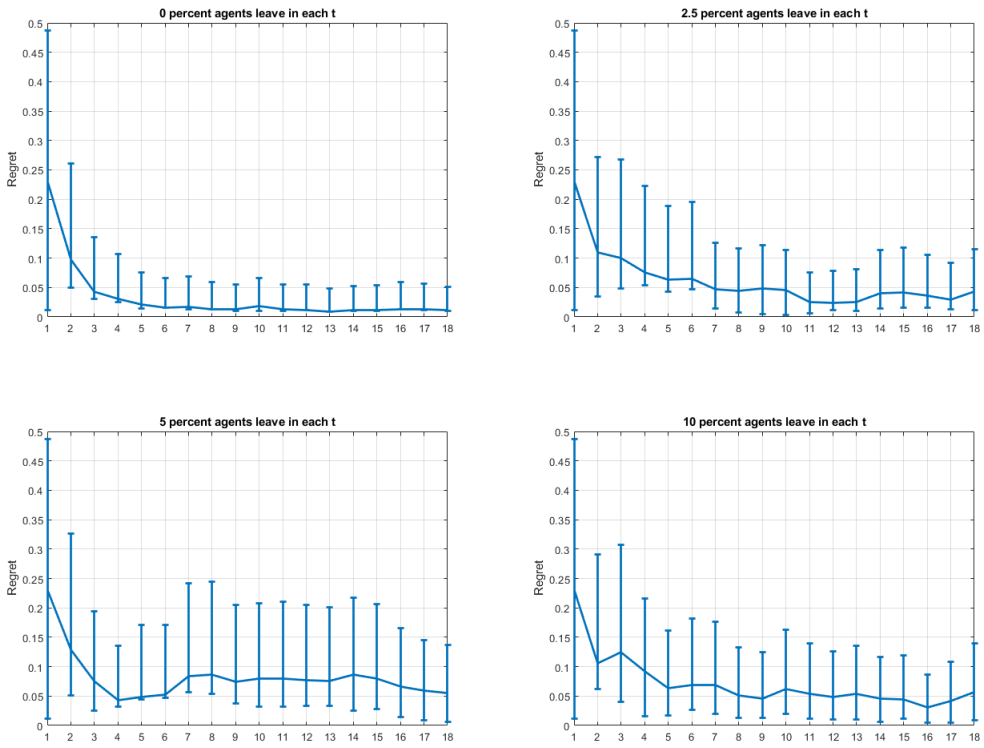


Figure 4: Average network production performance of our Hybrid Greedy-Thompson (HGT) algorithm in terms of regret (percentage) for four cases with different weekly worker turnover rates (0%, 2.5%, 5%, 10%). Standard deviations are calculated over 100 repetitions.

7 Conclusion

In conclusion, this paper develops a novel adaptive network formation algorithm tailored for settings where only edge-level outcomes are measured, and individual contributions are unobservable. We frame network formation as a Batched Multi-Armed Bandit (BMAB) problem and design a solution that leverages latent types within a flexible network model based on a variant of the stochastic block model, which can capture heterogeneous complementarities in a flexible way. By applying mean-field variational approximation, we obtain reliable batch-wise estimates of network production parameters and agents' latent types, accommodating a wide range of unobservable individual

differences.

Our adaptive network formation algorithm iteratively pairs agents through Bayesian updates and an optimized pairing policy, aiming to maximize production outcome and manage constraints like fair workload distribution. Simulation results indicate that our approach can effectively enhance productivity within limited batches and refine estimates of the production parameters and latent types. This iterative process demonstrates that the adaptive online framework we propose can efficiently navigate the trade-offs between exploration and exploitation, providing a robust strategy for network formation that can be applied across various organizational settings.

References

Agarwal, N. (2015), "An empirical model of the medical match," *American Economic Review*, 105, 1939–1978.

Ahmadpoor, M. and Jones, B. F. (2019), "Decoding team and individual impact in science and invention," *Proceedings of the National Academy of Sciences*, 116, 13885–13890.

Amini, A. A., Chen, A., Bickel, P. J., and Levina, E. (2013), "Pseudo-likelihood methods for community detection in large sparse networks," .

Bergeron, A., Bessone, P., Kabeya, J. K., Tourek, G., and Weigel, J. (2022), "Optimal Assignment of Bureaucrats: Evidence from Randomly Assigned Tax Collectors in the Drc," *NBER Working Paper*.

Bickel, P., Choi, D., Chang, X., and Zhang, H. (2013), "Asymptotic normality of maximum likelihood and its variational approximation for stochastic blockmodels," .

Bickel, P. J. and Chen, A. (2009), "A nonparametric view of network models and Newman–Girvan and other modularities," *Proceedings of the National Academy of Sciences*, 106, 21068–21073.

Bonhomme, S. (2021), "Teams: Heterogeneity, sorting, and complementarity," *arXiv preprint arXiv:2102.01802*.

Daudin, J.-J., Picard, F., and Robin, S. (2008), "A mixture model for random graphs," *Statistics and computing*, 18, 173–183.

Eichhorn, M., Banerjee, S., and Kempe, D. (2022), "Online team formation under different synergies," in *International Conference on Web and Internet Economics*, Springer, pp. 78–95.

Graham, B. and De Paula, Á. (2020), *The econometric analysis of network data*, Academic Press.

Hájek, J. (1970), "A characterization of limiting distributions of regular estimates," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 14, 323–330.

Hirano, K. and Porter, J. R. (2025), "Asymptotic representations for sequential decisions, adaptive experiments, and batched bandits," *arXiv preprint arXiv:2302.03117*.

Holland, P. W., Laskey, K. B., and Leinhardt, S. (1983), "Stochastic blockmodels: First steps," *Social networks*, 5, 109–137.

Jochmans, K. (2024), "Nonparametric Identification And Estimation of Stochastic Block Models From Many Small Networks"," .

Johari, R., Kamble, V., and Kanoria, Y. (2021), "Matching while learning," *Operations Research*, 69, 655–681.

Johari, R., Kamble, V., Krishnaswamy, A. K., and Li, H. (2018a), "Exploration vs. exploitation in team formation," *arXiv preprint arXiv:1809.06937*.

— (2018b), "Exploration vs. Exploitation in Team Formation for Collaborative Work," in *Conference on Web and Internet Economics (WINE)*.

Le Cam, L. et al. (1972), "Limits of experiments," in *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, vol. 1, pp. 245–261.

Lee, S. (2016), "Effect of online dating on assortative mating: Evidence from South Korea," *Journal of Applied Econometrics*, 31, 1120–1139.

Nowicki, K. and Snijders, T. A. B. (2001), "Estimation and prediction for stochastic blockstructures," *Journal of the American statistical association*, 96, 1077–1087.

Thompson, W. R. (1933), "On the likelihood that one unknown probability exceeds another in view of the evidence of two samples," *Biometrika*, 25, 285–294.

Van der Vaart, A. W. (2000), *Asymptotic statistics*, vol. 3, Cambridge university press.

Weidmann, B. and Deming, D. J. (2021), "Team players: How social skills improve team performance," *Econometrica*, 89, 2637–2657.

Xu, Y., Hu, M., Chu, J., and Ching, A. T. (2024), "Heterogeneous Complementarity and Team Design: The Case of Real Estate Agents," *Marketing Science*.

Zhang, K., Janson, L., and Murphy, S. (2020), "Inference for batched bandits," *Advances in neural information processing systems*, 33, 9818–9829.

A Proofs

A.1 Proof of Proposition 1

Proof of Part i). The joint likelihood ratio can be decomposed as follow:

$$\begin{aligned}\Lambda_f^{(n,m)}(\boldsymbol{\nu},\boldsymbol{\vartheta}) &= \log \frac{f_{\boldsymbol{\omega}+\frac{\boldsymbol{\nu}}{\sqrt{n}},\boldsymbol{\theta}+\frac{\boldsymbol{\vartheta}}{\sqrt{m}}}(\boldsymbol{Y},\boldsymbol{Z})}{f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z})} \\ &= \sum_{i=1}^{n} \log \frac{\pi_{\omega+\frac{\nu}{\sqrt{n}}}(Z_i)}{\pi_{\omega}(Z_i)} + \sum_{j=1}^{m} \log \frac{p_{\theta+\frac{\vartheta}{\sqrt{m}}}(Y_j|Z_{i_1(j)},Z_{i_2(j)})}{p_{\theta}(Y_j|Z_{i_1(j)},Z_{i_2(j)})} \\ &\equiv \Lambda_{\boldsymbol{\omega}}^{(n)}(\boldsymbol{\nu}) + \Lambda_{\boldsymbol{\theta}}^{(m)}(\boldsymbol{\vartheta}).\end{aligned} \tag{A.1}$$

For the first term, using the fact that $\omega(a) = \log \frac{\pi(a)}{1-\sum_{l=1}^{K-1}\pi(l)}$ implies $\pi_\omega(a) = \frac{e^{\omega(a)}}{1+\sum_{l=1}^{K-1}e^{\omega(l)}}$ for $a = 1,\ldots,K-1$, we have

$$\begin{aligned}\Lambda_{\boldsymbol{\omega}}^{(n)}(\boldsymbol{\nu}) &= \sum_{i=1}^{n}\left[\left(\omega(Z_i) + \frac{\nu(Z_i)}{\sqrt{n}} - \omega(Z_i)\right) - \log\frac{1+\sum_{b=1}^{K-1}e^{\omega(b)+\frac{\nu(b)}{\sqrt{n}}}}{1+\sum_{b=1}^{K-1}e^{\omega(b)}}\right] \\ &= \sum_{a=1}^{K-1}\nu(a)\Delta_{\boldsymbol{\omega}}^{(n)}(a) - \frac{1}{2}\sum_{a=1}^{K-1}\sum_{b=1}^{K-1}\nu(a)\mathcal{Q}_{\boldsymbol{\omega}}^{(n)}(a,b)\nu(b),\end{aligned}$$

where

$$\Delta_{\boldsymbol{\omega}}^{(n)}(a) = \frac{\partial\Lambda_{\boldsymbol{\omega}}(\boldsymbol{\nu})}{\partial\nu(a)} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\mathbb{1}\{Z_i = a\} - \pi(a)\right),$$

$$\mathcal{Q}_{\boldsymbol{\omega}}^{(n)}(a,b) = \begin{cases}\frac{\partial^2\Lambda_{\boldsymbol{\omega}}(\boldsymbol{\nu})}{\partial\nu(a)\partial\nu(a)} = \pi(a)(1-\pi(a)), & \text{if } a = b, \\ \frac{\partial^2\Lambda_{\boldsymbol{\omega}}(\boldsymbol{\nu})}{\partial\nu(a)\partial\nu(b)} = \pi(a)\pi(b), & \text{if } a \neq b.\end{cases}$$

For the second term, we have

$$\begin{aligned}
\Lambda_{\boldsymbol{\theta}}^{(m)}(\boldsymbol{\vartheta}) &= \sum_{j=1}^{m} \log \frac{p_{\theta+\frac{\vartheta}{\sqrt{m}}}(Y_j|Z_{i_1(j)}, Z_{i_2(j)})}{p_\theta(Y_j|Z_{i_1(j)}, Z_{i_2(j)})} \\
&= \sum_{j=1}^{m} \mathbb{1}\{Z_{i_1(j)} = a, Z_{i_2(j)} = b\} \log \frac{p_{\theta_{ab}+\frac{\vartheta_{ab}}{\sqrt{m}}}(Y_j|a,b)}{p_{\theta_{ab}}(Y_j|a,b)} \\
&= \sum_{j=1}^{m} \mathbb{1}\{Z_{i_1(j)} = a, Z_{i_2(j)} = b\} \left(\vartheta_{ab}\Delta_\theta(a,b) + \frac{1}{2}\vartheta_{ab}^2 \mathcal{I}_\theta(a,b) + o_{\mathrm{P}}(1) \right) \\
&= \sum_{a=1}^{K}\sum_{b=1}^{K} \left(\vartheta_{ab}^2 \Delta_{\boldsymbol{\theta}}(a,b) - \frac{1}{2}\vartheta_{ab}^2 \mathcal{Q}_{\boldsymbol{\theta}}(a,b) \right) + o_{\mathrm{P}}(1)
\end{aligned}$$

where

$$\begin{aligned}
\Delta_{\boldsymbol{\theta}}^{(m)}(a,b) &= \frac{1}{m}\sum_{j=1}^{m} \mathbb{1}\{Z_{i_1(j)} = a, Z_{i_2(j)} = b\}\Delta_\theta(a,b), \\
\mathcal{Q}_{\boldsymbol{\theta}}^{(m)}(a,b) &= \frac{1}{m}\sum_{j=1}^{m} \mathbb{1}\{Z_{i_1(j)} = a, Z_{i_2(j)} = b\}\mathcal{I}_\theta(a,b).
\end{aligned}$$

The third equality follows directly from the conditional LAN condition in Assumption 3 (a). □

Proof of Part ii). This part follows from a Lindeberg central limit theorem. □

Proof of Part iii). In the complete model $f_{\boldsymbol{\omega},\boldsymbol{\theta}}$, the MLE for $\boldsymbol{\pi}$ is given by

$$\hat{\pi}(a) = \frac{1}{n}\sum_{i=1}^{n} (\mathbb{1}\{Z_i = a\})$$

and $\hat{\omega}(a) = \log \pi(a) - \log(1 - \sum_{l=1}^{K-1} \pi(l))$ for $a = 1, \ldots, K-1$. From standard exponential family theory (see Bickel et al. (2013, Lemma 1)), it follows that

$$\sqrt{n}(\hat{\boldsymbol{\omega}}_f - \boldsymbol{\omega}) \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\mathcal{Q}}_{\boldsymbol{\omega}}^{-1}\right). \tag{A.2}$$

Next, for $\boldsymbol{\theta}$, Assumption 3 (b) ensures the asymptotic normality on the MLE for each individual production parameter θ_{ab}. Combining with the LAN result established above, the maximum likelihood estimate based on $f_{\boldsymbol{\omega},\boldsymbol{\theta}}$ satisfies

$$\sqrt{m}(\hat{\boldsymbol{\theta}}_f - \boldsymbol{\theta}) \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\mathcal{Q}}_{\boldsymbol{\theta}}^{-1}\right). \tag{A.3}$$

For further details, see, for example, Section 7.4 of Van der Vaart (2000). □

A.2 Proof of Theorem 1

Our proof builds on techniques presented in Bickel et al. (2013), where a similar result was established in their Theorem 1 but for the classical stochastic block model. To this end, we introduce the concept of *likelihood modularity*, defined as

$$Q^{(n,m)}(\boldsymbol{Y}, \boldsymbol{Z}) = \sup_{\boldsymbol{\omega},\boldsymbol{\theta}} \log f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{Z}).$$

We assume that it can be expressed as

$$Q^{(n,m)}(\boldsymbol{Y}, \boldsymbol{e}) = n \sum_{a=1}^{K} \pi_a(\boldsymbol{e}) \log(\pi_a(\boldsymbol{e})) + mF\left(\frac{\boldsymbol{O}(\boldsymbol{Y}, \boldsymbol{e})}{m}, \frac{\boldsymbol{m}(\boldsymbol{e})}{m}\right),$$

for $\boldsymbol{e} \in [K]^n$, where

$$m_{ab}(\boldsymbol{e}) = \sum_{j=1}^{m} \mathbb{1}\{e_{i_1(j)} = a, e_{i_2(j)} = b\}, \quad n_a(\boldsymbol{e}) = \sum_{i=1}^{n} \mathbb{1}\{e_i = a\},$$
$$O^r_{ab}(\boldsymbol{Y}, \boldsymbol{e}) = \sum_{j=1}^{m} \mathbb{1}\{e_{i_1(j)} = a, e_{i_2(j)} = b\} Y^r_j, \quad \pi_a(\boldsymbol{e}) = \frac{n_a(\boldsymbol{e})}{n},$$

and where $\boldsymbol{O}$ and $\boldsymbol{m}$ denote the collections of O^r_{ab} and m_{ab}, respectively. We will impose mild regularity conditions on F as stated in Assumption 5.

Define the confusion matrix $\boldsymbol{R} \in [0,1]^{K \times K}$ by

$$[\boldsymbol{R}(\boldsymbol{Z}, \boldsymbol{e})](a, b) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\{Z_i = b, e_i = a\}.$$

Let $S^r_\theta(a,b) = \mathbb{E}_\theta(Y^r_j \mid Z_{i_1(j)} = a, Z_{i_2(j)} = b)$ denote the r-th moment of an a-b pair, and let $\boldsymbol{S}^r_{\boldsymbol{\theta}}$ represent the $K \times K$ matrix collecting these values, $S^r_\theta(a,b)$. Abbreviate $\boldsymbol{R}(\boldsymbol{Z}, \boldsymbol{e}) \boldsymbol{S}^r_{\boldsymbol{\theta}} \boldsymbol{R}(\boldsymbol{Z}, \boldsymbol{e})'$ as $(\boldsymbol{R}\boldsymbol{S}^r_{\boldsymbol{\theta}}\boldsymbol{R}')(\boldsymbol{e})$, we define

$$\boldsymbol{X}^r(\boldsymbol{e}) = m^{-1}\boldsymbol{O}^r(\boldsymbol{Y}, \boldsymbol{e}) - (\boldsymbol{R}\boldsymbol{S}^r_{\boldsymbol{\theta}}\boldsymbol{R}')(\boldsymbol{e}).$$

The subsequent conditions will formalize the assumptions required for our analysis.

Assumption 4. *We assume the following two inequalities hold:*

$$P\left(\max_{\boldsymbol{e}} \|\boldsymbol{X}^r(\boldsymbol{e})\|_\infty > \epsilon\right) \leq 2K^{n+2}\exp\left(-C_1\epsilon^2 m\right) \tag{A.4}$$

for $\epsilon \leq c_1$*; and*

$$P\left(\max_{\boldsymbol{e}:|\boldsymbol{e}-\boldsymbol{Z}|\leq s} \|\boldsymbol{X}^r(\boldsymbol{e}) - \boldsymbol{X}^r(\boldsymbol{Z})\|_\infty > \epsilon\right) \leq 2\binom{n}{s} K^{s+2}\exp\left(-C_2\frac{n}{s}\epsilon^2 m\right) \tag{A.5}$$

for $\epsilon \geq c_2\frac{s}{n}$*, where* c_1*,* c_2*,* C_1*, and* C_2 *are some constants depending on the production model* p_θ *and the raw moment order* r*.*

Assumption 5. *We further assume that*

(a) The function $\boldsymbol{e} \mapsto F((\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}\boldsymbol{R}')(\boldsymbol{e}), \boldsymbol{R}\boldsymbol{\psi}\boldsymbol{R}'(\boldsymbol{e}))$ *is maximized by any* $\boldsymbol{e} \in \mathcal{S}_{\boldsymbol{Z}}$*.*

(b) F *is Lipschitz in its arguments; The directional derivatives*

$$\frac{\partial^2 F}{\partial \epsilon^2}\left(\boldsymbol{A}_0 + \epsilon(\boldsymbol{A}_1 - \boldsymbol{A}_0), \boldsymbol{B}_0 + \epsilon(\boldsymbol{B}_1 - \boldsymbol{B}_0)\right)|_{\epsilon=0^+}$$

are continuous in $(\boldsymbol{A}_1, \boldsymbol{B}_1)$ *for all* $(\boldsymbol{A}_0, \boldsymbol{B}_0)$ *in a neighborhood of* $(diag(\pi)S_{\boldsymbol{\theta}}diag(\pi), diag(\pi))$*.*

(c) Let $G(\boldsymbol{R}, \boldsymbol{S}) = F(\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}\boldsymbol{R}', \boldsymbol{R}\boldsymbol{\psi}\boldsymbol{R}')$*, assume that*

$$\left.\frac{\partial G((1-\epsilon)diag(\pi) + \epsilon\boldsymbol{R}, \boldsymbol{S})}{\partial \epsilon}\right|_{\epsilon=0^+} < -C < 0,$$

for all $\boldsymbol{R} \in \{\boldsymbol{R} : \boldsymbol{R} \geq 0, \boldsymbol{R}'\boldsymbol{1} = \pi\}$*.*

Proof of Part (i). Following Bickel et al. (2013), we proceed by addressing two cases separately: when $\boldsymbol{e}$ is far from $\boldsymbol{Z}$ and when $\boldsymbol{e}$ is close to $\boldsymbol{Z}$. For this purpose, we define set

$$E_{\delta_n} = \left\{\boldsymbol{e} : |F((\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}^r\boldsymbol{R}')(\boldsymbol{e}), \boldsymbol{R}\boldsymbol{\psi}\boldsymbol{R}'(\boldsymbol{e})) - F((\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}^r\boldsymbol{R}')(\boldsymbol{Z}), \boldsymbol{R}\boldsymbol{\psi}\boldsymbol{R}'(\boldsymbol{Z}))| \geq \delta_n\right\}$$

where δ_n is a sequence slowly converges to zero. By (A.4), which implies that $m^{-1}\boldsymbol{O}^r(\boldsymbol{Y}, \boldsymbol{e}) \to (\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}^r\boldsymbol{R}')(\boldsymbol{e})$, and the continuity of F, we have

$$P\left(\max_{\boldsymbol{e}}\left|F\left(\frac{\boldsymbol{O}^r(\boldsymbol{Y}, \boldsymbol{e})}{m}, \frac{\boldsymbol{m}(\boldsymbol{e})}{m}\right) - F((\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}^r\boldsymbol{R}')(\boldsymbol{e}), \boldsymbol{R}\boldsymbol{\psi}\boldsymbol{R}'(\boldsymbol{e}))\right| \geq \frac{\delta_n}{2}\right) = o(1).$$

Therefore, for all $\boldsymbol{e} \in E_{\delta_n}$, we have

$$F\left(\frac{\boldsymbol{O}^r(\boldsymbol{Y},\boldsymbol{e})}{m}, \frac{\boldsymbol{m}(\boldsymbol{e})}{m}\right) \leq F((\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}^r\boldsymbol{R}')(\boldsymbol{Z}), \boldsymbol{R}\boldsymbol{\psi}\boldsymbol{R}'(\boldsymbol{Z})) - \frac{\delta_n}{2} + o_P(\delta_n),$$

and subsequently,

$$\begin{aligned}
&\sum_{\boldsymbol{e}\in E_{\delta_n}} e^{\sup_{\boldsymbol{\omega},\boldsymbol{\theta}} \log f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{e})} \\
&= \sum_{\boldsymbol{e}\in E_{\delta_n}} e^{n\sum_{a=1}^K \pi_a(\boldsymbol{e})\log(\pi_a(\boldsymbol{e})) + mF((\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}^r\boldsymbol{R}')(\boldsymbol{e}), \boldsymbol{R}\boldsymbol{\psi}\boldsymbol{R}'(\boldsymbol{e}))} \\
&= \sum_{\boldsymbol{e}\in E_{\delta_n}} e^{O(n)+m(F((\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}^r\boldsymbol{R}')(\boldsymbol{Z}), \boldsymbol{R}\boldsymbol{\psi}\boldsymbol{R}'(\boldsymbol{Z})) - \frac{\delta_n}{2} + o_P(\delta_n))} \\
&= e^{mF((\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}^r\boldsymbol{R}')(\boldsymbol{Z}), \boldsymbol{R}\boldsymbol{\psi}\boldsymbol{R}'(\boldsymbol{Z}))} e^{O(n) - \frac{m\delta_n}{2} + o_P(m\delta_n)} K^n \\
&= e^{\sup_{\boldsymbol{\omega},\boldsymbol{\theta}} \log f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z})} o_P(1)
\end{aligned} \tag{A.6}$$

where the last equality holds if we choose δ_n slow enough such that $m\delta_n \gg n$.

We now consider the case where $\boldsymbol{e} \notin E_{\delta_n}$, meaning that $\boldsymbol{e}$ is close to $\boldsymbol{Z}$. Let $|\boldsymbol{e} - \boldsymbol{Z}|$ denote $\sum_{i=1}^n \mathbb{1}\{\boldsymbol{e}_i \neq \boldsymbol{Z}_i\}$, and let $\mathcal{S}(\boldsymbol{Z})$ denote the set of equivalent labels $\boldsymbol{Z}'$ up to permutation $\tau : [K] \to [K]$, that is, $\boldsymbol{Z}' = \tau(\boldsymbol{Z})$ for some τ. For a given $\boldsymbol{e}$, define $\bar{\boldsymbol{e}} = \arg\min_\tau |\tau(\boldsymbol{e}) - \boldsymbol{Z}|$. From Eq. (A.5), we have

$$\begin{aligned}
&P\left(\max_{\boldsymbol{e}\notin\mathcal{S}(\boldsymbol{Z})} \|\boldsymbol{X}^r(\bar{\boldsymbol{e}}) - \boldsymbol{X}^r(\boldsymbol{Z})\|_\infty > \epsilon\frac{|\bar{\boldsymbol{e}} - \boldsymbol{Z}|}{n}\right) \\
&\leq \sum_{s=1}^n P\left(\max_{\boldsymbol{e}:|\bar{\boldsymbol{e}}-\boldsymbol{Z}|=s} \|\boldsymbol{X}^r(\bar{\boldsymbol{e}}) - \boldsymbol{X}^r(\boldsymbol{Z})\|_\infty > \epsilon\frac{s}{n}\right) \\
&\leq \sum_{s=1}^n 2K^K n^s K^{s+2} \exp\left(-Cm\frac{s}{n}\epsilon^2\right) \to 0,
\end{aligned}$$

where the second inequality follows from (A.5). The final convergence to zero is due to the condition $m \gg n\log n$, which ensures that we can choose $\epsilon \to 0$ such that $m\epsilon^2 \gg n\log n$. This result establishes that

$$\max_{\boldsymbol{e}\notin\mathcal{S}(\boldsymbol{Z})} \|\boldsymbol{X}^r(\bar{\boldsymbol{e}}) - \boldsymbol{X}^r(\boldsymbol{Z})\|_\infty = o_P(|\bar{\boldsymbol{e}} - \boldsymbol{Z}|/n).$$

As a result, for $\boldsymbol{e} \notin E_{\delta_n}$, we have

$$\left\|\frac{\boldsymbol{O}^r(\boldsymbol{Y},\bar{\boldsymbol{e}})}{m} - \frac{\boldsymbol{O}^r(\boldsymbol{Y},\boldsymbol{Z})}{m}\right\|_\infty = \frac{|\bar{\boldsymbol{e}} - \boldsymbol{Z}|}{n} o_P(1) + \left\|(\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}^r\boldsymbol{R}')(\bar{\boldsymbol{e}}) - (\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}^r\boldsymbol{R}')(\boldsymbol{Z})\right\|_\infty.$$

This further implies

$$\frac{\boldsymbol{O}^r(\boldsymbol{Y},\bar{\boldsymbol{e}})}{m} - \frac{\boldsymbol{O}^r(\boldsymbol{Y},\boldsymbol{Z})}{m} = \left((\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}^r\boldsymbol{R}')(\bar{\boldsymbol{e}}) - (\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}^r\boldsymbol{R}')(\boldsymbol{Z})\right)(1+o_P(1)), \tag{A.7}$$

where $o_P(1)$ is uniform over $\boldsymbol{e}$.

From Assumption 5 (c), we know that

$$\frac{\partial}{\partial\epsilon}F\left((1-\epsilon)(\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}^r\boldsymbol{R}')(\boldsymbol{Z}) + \epsilon\boldsymbol{A}, (1-\epsilon)(\boldsymbol{R}\psi\boldsymbol{R}')(\boldsymbol{Z}) + \epsilon\boldsymbol{B}\right)\Big|_{\epsilon=0^+} < -\Omega_P(1)$$

for $(\boldsymbol{A},\boldsymbol{B})$ lies in the neighborhood of $((\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}^r\boldsymbol{R}')(\boldsymbol{Z}),(\boldsymbol{R}\psi\boldsymbol{R}')(\boldsymbol{Z}))$, where $\Omega_P(a_n)$ denotes that a_n is bounded below in probability by b_n (up to multiplying a constant factor). Using the result in (A.7), alone with the facts that $\|\boldsymbol{O}^r(\boldsymbol{Y},\boldsymbol{Z})/m - (\boldsymbol{R}\boldsymbol{S}_{\boldsymbol{\theta}}^r\boldsymbol{R}')(\boldsymbol{Z})\| = o_P(1)$ and $\|\boldsymbol{m}(\boldsymbol{Z})/m - (\boldsymbol{R}\psi\boldsymbol{R}')(\boldsymbol{Z})\| = o_P(1)$ (as implied by (A.4)), we can apply the steps outlined in Bickel et al. (2013, page 19) to show that

$$\frac{\partial}{\partial\epsilon}F\left((1-\epsilon)\frac{\boldsymbol{O}^r(\boldsymbol{Y},\boldsymbol{Z})}{m} + \epsilon\boldsymbol{A}, (1-\epsilon)\frac{\boldsymbol{m}(\boldsymbol{Z})}{m} + \epsilon\boldsymbol{B}\right)\Big|_{\epsilon=0^+} < -\Omega_P(1)$$

for $(\boldsymbol{A},\boldsymbol{B})$ in the neighborhood of $(\frac{\boldsymbol{O}^r(\boldsymbol{Y},\boldsymbol{Z})}{m},\frac{\boldsymbol{m}(\boldsymbol{Z})}{m})$. Since $\boldsymbol{e}\notin E_{\delta_n}$ and $\delta_n\to 0$, it follows that

$$F\left(\frac{\boldsymbol{O}^r(\boldsymbol{Y},\boldsymbol{e})}{m},\frac{\boldsymbol{m}(\boldsymbol{e})}{m}\right) - F\left(\frac{\boldsymbol{O}^r(\boldsymbol{Y},\boldsymbol{Z})}{m},\frac{\boldsymbol{m}(\boldsymbol{Z})}{m}\right) \le -\frac{1}{n}\Omega_P(|\bar{\boldsymbol{e}}-\boldsymbol{Z}|), \tag{A.8}$$

where $\Omega_P(|\bar{\boldsymbol{e}}-\boldsymbol{Z}|)$ is uniform over $\boldsymbol{e}$.

Therefore, for $|\bar{\boldsymbol{e}}-\boldsymbol{Z}| = s$, we have

$$\begin{aligned}
&\sup_{\boldsymbol{\omega},\boldsymbol{\theta}}\log f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{e}) - \sup_{\boldsymbol{\omega},\boldsymbol{\theta}}\log f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z})\\
\le\ & n\left|\sum_{a=1}^{K}(\pi_a(\boldsymbol{e})\log(\pi_a(\boldsymbol{e})) - \pi_a(\boldsymbol{Z})\log(\pi_a(\boldsymbol{Z})))\right|\\
&+ m\left(F\left(\frac{\boldsymbol{O}(\boldsymbol{Y},\boldsymbol{e})}{m},\frac{\boldsymbol{m}(\boldsymbol{e})}{m}\right) - F\left(\frac{\boldsymbol{O}(\boldsymbol{Y},\boldsymbol{Z})}{m},\frac{\boldsymbol{m}(\boldsymbol{Z})}{m}\right)\right)\\
\le\ & nO(1) - m\Omega_P(s/n).
\end{aligned}$$

Thus, we obtain

$$
\begin{aligned}
&\sum_{\boldsymbol{e}\notin E_{\delta_n},\boldsymbol{e}\notin\mathcal{S}(\boldsymbol{Z})} e^{\sup_{\boldsymbol{\omega},\boldsymbol{\theta}}\log f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{e})} \\
&= \sum_{s=1}^{n}\sum_{\boldsymbol{e}:|\bar{\boldsymbol{e}}-\boldsymbol{Z}|=s} e^{\sup_{\boldsymbol{\omega},\boldsymbol{\theta}}\log f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{e})} \\
&= \sum_{s=1}^{n}\sum_{\boldsymbol{e}:|\bar{\boldsymbol{e}}-\boldsymbol{Z}|=s} e^{\sup_{\boldsymbol{\omega},\boldsymbol{\theta}}\log f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z})} e^{\sup_{\boldsymbol{\omega},\boldsymbol{\theta}}\log f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{e})-\sup_{\boldsymbol{\omega},\boldsymbol{\theta}}\log f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z})} \\
&\le \sum_{s=1}^{n}\sum_{\boldsymbol{e}:|\bar{\boldsymbol{e}}-\boldsymbol{Z}|=s} e^{\sup_{\boldsymbol{\omega},\boldsymbol{\theta}}\log f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z})} e^{O(n)-m\Omega_P(s/n)} \\
&\le \sum_{s=1}^{n} e^{\sup_{\boldsymbol{\omega},\boldsymbol{\theta}}\log f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z})} K^K n^s K^s e^{O(n)-m\Omega_P(s/n)} \\
&\le \sum_{s=1}^{n} e^{\sup_{\boldsymbol{\omega},\boldsymbol{\theta}}\log f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z})} K^K e^{s(logn+\log K+O(n)-\Omega_P(m/n))} \\
&= e^{\sup_{\boldsymbol{\omega},\boldsymbol{\theta}}\log f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z})} o_P(1).
\end{aligned}
\tag{A.9}
$$

Combining this result with (A.6), we conclude that for all $(\boldsymbol{\omega},\boldsymbol{\theta})$,

$$
\sum_{\boldsymbol{e}\notin\mathcal{S}(\boldsymbol{Z})} f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{e}) \le \sum_{\boldsymbol{e}\notin\mathcal{S}(\boldsymbol{Z})} \sup_{\boldsymbol{\omega},\boldsymbol{\theta}} f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{e}) = \sup_{\boldsymbol{\omega},\boldsymbol{\theta}} f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z}) o_P(1). \tag{A.10}
$$

Recalling that $g_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}) = \sum_{\boldsymbol{e}\in[K]^n} f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{e})$, we obtain the following inequality for all $(\boldsymbol{\omega},\boldsymbol{\theta})$,

$$
\begin{aligned}
f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z}) \le g_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}) &= f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z}) + \sum_{\boldsymbol{e}\notin\mathcal{S}(\boldsymbol{Z})} f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{e}) \\
&\le f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z}) + \sup_{\boldsymbol{\omega},\boldsymbol{\theta}} f_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z}) o_P(1) \\
&= f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{Z}) + o_P(1),
\end{aligned}
\tag{A.11}
$$

where the first inequality follows from the non-negativity of $f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y},\boldsymbol{e})$ for any $\boldsymbol{e}$, and the second inequality is from (A.10).

□

Proof of Part (ii). Since the MLE $(\hat{\boldsymbol{\omega}}_g,\hat{\boldsymbol{\theta}}_g)$ is obtained by maximizing $g_{\boldsymbol{\omega},\boldsymbol{\theta}}(\boldsymbol{Y})$, it follows from Part (i) that there exist an element $(\hat{\boldsymbol{\omega}}'_g,\hat{\boldsymbol{\theta}}'_g) \in \mathcal{S}(\hat{\boldsymbol{\omega}}_g,\hat{\boldsymbol{\theta}}_g)$ such that

$$
|f_{\boldsymbol{\pi}'_g,\boldsymbol{\theta}'_g}(\boldsymbol{Y},\boldsymbol{Z}) - f_{\hat{\boldsymbol{\pi}}_f,\hat{\boldsymbol{\theta}}_f}(\boldsymbol{Y},\boldsymbol{Z})| = o_P(1), \tag{A.12}
$$

where $(\hat{\boldsymbol{\pi}}_f, \hat{\boldsymbol{\theta}}_f)$ is the MLE that maximizes $f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{Z})$. By the LAN result established in Proposition 1 for the complete model f, we then have

$$\frac{1}{\sqrt{n}}|\hat{\boldsymbol{\omega}}'_g - \hat{\boldsymbol{\omega}}_f| = o_P(1) \quad \text{and} \quad \frac{1}{\sqrt{m}}|\hat{\boldsymbol{\theta}}'_g - \hat{\boldsymbol{\theta}}_f| = o_P(1). \tag{A.13}$$

This can be seen more clearly by contradiction: if either equality in (A.13) does not hold, then (A.12) is violated.

Combining (A.2), (A.3) with (A.13) completes the proof. □

A.3 Proof of Theorem 2

Proof of Part (i). Recall from (3.6) that

$$\begin{aligned} J_{\boldsymbol{\pi},\boldsymbol{\theta};\boldsymbol{q}}(\boldsymbol{Y}) = & \sum_{i=1}^{n}\sum_{a=1}^{K} q_i(a)\left(-\log q_i(a) + \log \pi(a)\right) \\ & + \sum_{j=1}^{m}\sum_{a=1}^{K}\sum_{b=1}^{K} q_{i_1(j)}(a) q_{i_2(j)}(b) \log p_\theta(Y_j|a,b). \end{aligned}$$

It is straightforward to verify that $J_{\boldsymbol{\pi},\boldsymbol{\theta};\boldsymbol{q}}(\boldsymbol{Y}) = \log f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{Z})$ when $q_i(a) = \mathbb{1}\{Z_i = a\}$ for all $i = 1, \ldots, n$ and $a \in [K]$. Therefore, let $\hat{\boldsymbol{q}} = \hat{\boldsymbol{q}}(\boldsymbol{\pi}, \boldsymbol{\theta}) \equiv \arg\max_{\boldsymbol{q}} J_{\boldsymbol{\pi},\boldsymbol{\theta};\boldsymbol{q}}(\boldsymbol{Y})$, it follows that $\log f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{Z}) \le J_{\boldsymbol{\pi},\boldsymbol{\theta};\hat{\boldsymbol{q}}}(\boldsymbol{Y})$.

On the other hand, by (3.5) and the non-negativity of the KL divergence, we have $J_{\boldsymbol{\pi},\boldsymbol{\theta};\boldsymbol{q}}(\boldsymbol{Y}) \le \log g_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y})$ for all q. Combining these results, we arrive at

$$\log f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{Z}) \le J_{\boldsymbol{\pi},\boldsymbol{\theta};\hat{\boldsymbol{q}}}(\boldsymbol{Y}) \le \log g_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}). \tag{A.14}$$

Furthermore, we observe that

$$\begin{aligned} \log g_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}) &= \log\left(f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{Z}) + \sum_{\boldsymbol{e}\notin\mathcal{S}(\boldsymbol{Z})} f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{e})\right) \\ &= \log f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{Z}) + \log\left(1 + \frac{\sum_{\boldsymbol{e}\notin\mathcal{S}(\boldsymbol{Z})} f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{e})}{f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{Z})}\right) \\ &= \log f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{Z}) + \log\left(1 + \frac{f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{Z}) o_P(1)}{f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{Z})}\right) \\ &= \log f_{\boldsymbol{\pi},\boldsymbol{\theta}}(\boldsymbol{Y}, \boldsymbol{Z}) + o_P(1), \end{aligned}$$

where the third equality follows from (A.11) in the proof of Theorem 1. This result, together with (A.14), completes the proof. □

Proof of Part (ii). The proof follows using arguments similar to Theorem 1 (ii). □