A Unified Adaptive Testing System Enabled by Hierarchical Structure Search

Junhao Yu [1] **Yan Zhuang** [1] **Zhenya Huang** [1 2] **Qi Liu** [1 2] **Xin Li** [1] **Rui Li** [3] **Enhong Chen** [1]

Abstract

Adaptive Testing System (ATS) is a promising testing mode, extensively utilized in standardized tests like the GRE. It offers a personalized ability assessment by dynamically adjusting questions based on individual ability levels. Compared to traditional exams, ATS can improve the accuracy of ability estimates while simultaneously reducing the number of questions required. Despite the diverse ATS testing formats, tailored to different adaptability requirements in various testing scenarios, there is a notable absence of a unified framework to model them. In this paper, we introduce a unified data-driven ATS framework that conceptualizes the various testing formats as a hierarchical test structure search problem. It can learn directly from data to solve optimal questions for each student, eliminating the need for manual test design. The proposed solution algorithm comes with theoretical guarantees for the estimation error and convergence. Empirical results show that our framework maintains assessment accuracy while reducing question count by 20% on average and improving training stability.

1. Introduction

Adaptive Testing System (ATS) is a pioneering application of intelligent education, providing a personalized and efficient assessment experience (Chen et al., 2015; Vie et al., 2017). Moving beyond the traditional paper-and-pencil tests, which uniformly present the same questions to all examinees, ATS dynamically adapts the questions to the individual's ability level. This approach not only enhances the precision of the assessment but also reduces the number of questions required, thereby enhancing testing efficiency. The adoption of ATS in high-stakes exams such as the Duolingo Test, GRE, TOEFL, and GMAT.

[1]State Key Laboratory of Cognitive Intelligence, University of Science and Technology of China [2]Institute of Artificial Intelligence, Hefei Comprehensive National Science Center [3]School of Computer Science and Technology, Xidian University. Correspondence to: Enhong Chen <cheneh@ustc.edu.cn>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024.

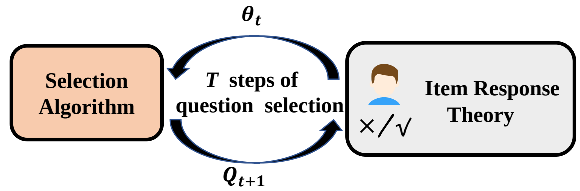


Figure 1. A traditional process of ATS: At step t, IRT estimates the current ability θ_t based on the previous t steps of records. Then the selection algorithm selects the next set of questions Q_{t+1} based on the current estimated ability θ_t.

ATS is an interactive process between students and the system. As illustrated in Figure 1, at each test step, cognitive diagnosis models(Wang et al., 2023a) like Item Response Theory (IRT) (Embretson & Reise, 2013), a user model based on psychometrics, first estimates the student's ability by analyzing previous responses. Then, the Selection Algorithm selects the next questions from the pool based on Reinforcement Learning (Zhuang et al., 2022) or specific informativeness metrics (Chang & Ying, 1996; Bi et al., 2020), which typically aim to match question difficulty with the student's estimated ability (Lord, 2012). This iterative cycle continues until a stopping rule is satisfied. ATS's core is to optimize accuracy in estimating student ability with the fewest questions necessary.

Current ATS can be categorized into two research domains based on the test format: Computerized Adaptive Testing (CAT) and Multistage Testing (MST). A recent comprehensive survey (Liu et al., 2024) thoroughly discussed these two approaches: CAT selects one question at each step, requiring students to complete one question before progressing to the next (Van der Linden & Pashley, 2009), whereas MST selects a set of questions (a testlet) at each step. These testlets are predefined by experts and typically contain questions of similar difficulty and a diverse range of content (Sari et al., 2016). CAT is known for its precise estimation due to its granular approach, whereas MST can incorporate expert insights to offer a more comprehensive testing experience (Magis et al., 2017). *Despite both serving the core objective of adaptive assessment, they currently function independently due to their different granularities in adaptivity.* CAT at question level and MST at testlet level.

There is a clear need for a unified framework that can flexibly model both CAT and MST to harmonize these methodologies within a single ATS approach. Such a unified framework would benefit from the high adaptability of CAT while also embracing the combinatorial effects present within MST's testlets. Crucially, it could avoid the time-consuming and labor-intensive process of expert testlet design. However, developing such a unified framework is not trivial. For instance: What mechanisms can be employed to determine the quantity of questions to select at each step, without manual intervention? How to automatically optimize the desired test forms/structures directly from the available data? How should we account for the similarity and combinatorial effects among the questions within these auto-generated testlets? These challenges underscore the complexity of the decision-making problem at hand.

This paper introduces a data-driven framework that treats ATS as a *structure search problem*. We envision each potential question as a node within a network graph, with our objective being to search the most effective pathways—or edges—in the graph for individual students. It enables a systematic and automated approach to question selection or testlet construction. Our contributions are as follows:

- We redefine various forms of ATS as a unified problem of test structure search for the first time. It is data-driven and dynamically navigates question selection, eliminating the need for manual expert design.
- A differentiable and hierarchical optimization algorithm is designed to select appropriate questions. This approach enables efficient and accurate optimization with theoretical guarantees of convergence and gradient estimation error.
- This framework exhibits superior performance on various real-world datasets, surpassing other ATS methods in accuracy and efficiency. Notably, it exhibits more stable convergence during training compared to other data-driven approaches.

2. Background and Related Works

Adaptive Testing System (ATS) is an iterative process, it mainly consists of two components: Item Response Theory and question selection algorithm. Below, we will introduce these two components separately:

(1) Item Response Theory (IRT). IRT is grounded in psychometrics and cognitive science (Ackerman et al., 2003; Huang et al., 2020), adopting question's and student's features to predict the response (correct or wrong). The simplest form is the one-parameter logistic (1PL) model:

$$\Pr(\text{student answers question } j \text{ correctly}) = \sigma(\theta - \beta_j),$$

where $\sigma(\cdot)$ is the logistic function, $\beta \in \mathbb{R}$ represents each question's pre-calibrated parameter called *difficulty*, and $\theta \in \mathbb{R}$ is student's latent ability to be estimated. Recently, there have been many studies incorporating neural networks to model student-question interactions .(Cheng et al., 2019; Gao et al., 2022; Wang et al., 2023a; Shen et al., 2024)

(2) Selection Algorithms. The purpose of ATS is to estimate student ability with the fewest questions while maintaining accuracy. Therefore, the selection algorithm is a crucial step. Traditional ATS have approached this issue from two perspectives, adaptability and flexibility. CAT focuses more on adaptability, aiming to get an optimal test (Meijer & Nering, 1999). It selects only one question at one step. Various algorithms on information-based selection (Lord, 2012; Chang & Ying, 1996; Rudner, 2002; van der Linden, 1998; Veerkamp & Berger, 1997; Kang et al., 2017; Ma et al., 2023), data-driven approaches (Nurakhmetov, 2019; Zhuang et al., 2022; Ghosh & Lan, 2021; Wang et al., 2023b; Li et al., 2023; Yu et al., 2023) and other approaches(Veldkamp & Verschoor, 2019; Gilavert & Freire, 2022; Feng et al., 2023; Mujtaba & Mahapatra, 2021) have been proposed. However, such algorithms impose limitations on students, as selecting one question at one time restricts their question-solving habits (Mead, 2006). In this regard, MST can address this problem by recommending a set of question at once, greatly enhancing the flexibility (Sari et al., 2016). Students can freely choose questions at one step. Nevertheless, this approach has drawbacks as it requires more questions to be completed and demands expert design of question sets (Magis et al., 2017). Currently, there is a lack of a comprehensive approach that can uniformly address the adaptability and flexibility issue.

In this paper, we propose a unified framework for ATS called UATS that addresses the respective challenges encountered in CAT and MST. The framework encompasses both CAT and MST, facilitating automatic composition of question packages during the testing process. Our findings reveal that the UATS framework exhibits exceptional performance in both theoretical analysis and experimental evaluations, surpassing the capabilities of CAT and MST approaches.

3. Task Formalization

The objective of an Adaptive Testing System (ATS) is to sequentially present a student with an optimally tailored sequence of questions or testlets. In essence, at each step of the process, the most appropriate set of questions for the student, $Q_t = \{q_t^1, q_t^2, ..., q_t^m\}$, is selected for administration.

For CAT, the set size m is 1, indicating a single question q is selected at a time, while for MST, a set of questions or a testlet of size m is chosen. The ultimate goal of ATS are twofold: **(1)** select valuable and best-fitting questions for each student, minimizing the length of the test; **(2)** utilize the student's responses to accurately estimate their true ability when the test is over[1].

To achieve these goals, an ATS comprises two key components: the Item Response Theory (IRT) and the selection algorithm. As depicted in Figure 1, at each test step , the ATS first employs IRT to estimate the student's current ability θ_t based on responses to the previous t question sets . The binary cross-entropy loss is frequently utilized: given responses to previous t sets: $D_t = \{(Q_1, Y_1), \ldots, (Q_t, Y_t)\}$, the empirical loss is

$$L(\theta, D_t) = \sum_{i=1}^{t} -\log p_\theta(Q_i, Y_i), \tag{1}$$

where $Y_i = \{y_i^1, y_i^2, ..., y_i^m\}$ represents student's response labels to Q_i and $y^m = 1$ indicates a correct response to question q^m, and 0 otherwise; $p_\theta(Q, Y)$ represents the probability of the response (Q, Y) towards a student with θ, and its specific form is determined by IRT. Thus the current estimate of ability, θ_t, is obtained by minimizing the loss function $L(\theta, D_t)$: $\theta_t = \arg\min_\theta L(\theta, D_t)$.

Then, the selection algorithm picks the next question set Q_{t+1} from the question bank, using the student's current ability estimate θ_t as a guide: $Q_{t+1} \sim \pi(\theta_t)$, where π can be various criteria that measure how much information the question set will provide about the student's ability (Lord, 2012; Chang & Ying, 1996), or could be the output of a policy trained by data-driven methods (e.g., Reinforcement Learning) (Zhuang et al., 2022; Ghosh & Lan, 2021).

Evaluation Methods: The ground truth of a student's true ability is unavailable in datasets, presenting a challenge to evaluate the ability estimate. To address this, a common practice is to randomly divide the data of each student into a query set D_u and a support set D_t (Ghosh & Lan, 2021). The support set D_t is used to simulate question selection process and estimate the final ability value θ_T, while the query set D_u is employed to assess the accuracy of this estimates. Cross-entropy $L(\theta_T, D_u)$ or various classification metrics such as prediction accuracy can be employed for evaluation. The details can be found in Section 5.

4. Test Structure Search

In traditional ATS introduced above, the question selection algorithm determines the most appropriate set of questions,

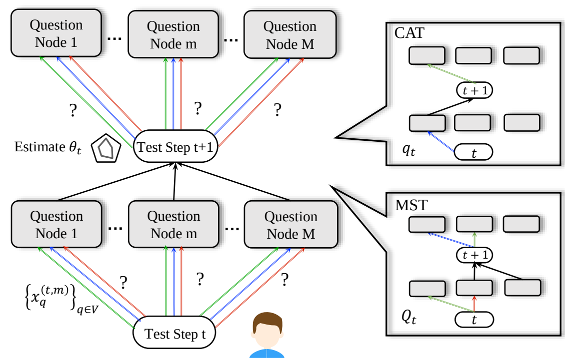


Figure 2. The left structure demonstrates our UATS framework. Each layer selects questions from the M question pools, and each edge corresponds to questions from a question node. The selection of questions is transformed into a structural search problem. The right illustration illustrates the implementation of CAT and MST in our scenario.CAT selects only one question at a time, while MST selects a group of questions at a time.

denoted as Q_t, for the student to answer at each step t, based on the student's current ability estimate θ_t. However, due to the diverse nature of testing scenarios, the number of questions selected at each step may vary. This variability poses a challenge as it requires considering various factors that influence the test, such as question similarity and combination effects. This significantly increases the complexity of the selection algorithm and cannot be easily achieved through expert intervention, which involves predefining sets of candidate question sets (testlets). This manual approach is time-consuming, labor-intensive and cannot adapt to more complex testing scenarios.

To address this, this paper proposes a data-driven approach that models different testing forms as a unified structure search problem. Specifically, from a global perspective, *adaptive testing can be viewed as a test structure search problem from the initial node to the terminal node*. The left part of Figure 2 illustrates this concept. The number of question nodes M determines the maximum number of questions selected at each step, and different edge colors between nodes represent different questions. In the case of CAT shown in the Figure 2 on the right, only one question is selected at each step, resulting in a single edge between test steps. The other one is the case of MST. M questions is selected, leading to M edges between test nodes.

4.1. Problem Definition

Each directed edge from test node t to question node m is represented by $\{x_q^{(t,m)}\}_{q\in V}$ and $x_q^{(t,m)} \in \{0, 1\}$, where 1 indicates that the edge is selected in the final test structure, and 0 otherwise. For each pair of nodes (t, m), there are a total of $|V| + 1$ edges , including all the questions in the

[1]An important assumption (Chang, 2015) of ATS is that student's true ability level θ^* is constant throughout the test.

question bank V and an additional zero edge (representing no selection). Only one edge is chosen for the final structure. To optimize the test structure directly from large-scale student response data, each student's response data are further randomly divided into a query set D_u and a support set D_t. The ability estimation loss function is used to optimize the test structure based on the query set D_u, while the support set D_t is utilized to optimize the optimal ability estimate given the structure parameters:

$$\min_{x} L(\theta_t^*(x), D_u) \tag{2}$$

$$\text{s.t. } \theta_t^*(x) = \arg\min_{\theta} L(\theta, D_t(x)). \tag{3}$$

where $\theta_t(x)$ represents student's final ability estimate at based on their response on the questions corresponding to the structural parameter x. Therefore, at each step t, the selected question q_t^m is determined as follows:

$$q_t^m = \arg\max_{q \in V} x_q^{(t,m)} \tag{4}$$

Ultimately, the questions that correspond to the structural parameter x being equal to 1 are chosen to form the final set of selected questions, denoted as $Q_t = \{q_t^1, q_t^2, ..., q_t^M\}$.

Obviously, structural parameter x is discrete and non-differentiable. We further make a relaxation of the discrete structural parameters: a set of weight parameters, denoted as $\{\alpha_q^{(t,m)}\}_{q \in V}$, is defined between node t and node m. This leads to the unified definition of ATS, referred to as the Test Structure Search problem:

Definition 4.1 (Unified Definition of ATS)**.** Different forms of Adaptive Testing System can be unified as a Test Structure Search problem. A set of weight parameters $\{\alpha_q^{(t,m)}\}_{q \in V}$, related to the structure is defined to achieve the differentiable optimization:

$$\min_{\alpha} L(\theta_t^*(\alpha), D_u) \tag{5}$$

$$\text{s.t. } \theta_t^*(\alpha) = \arg\min_{\theta} L(\theta, D_t(\alpha)). \tag{6}$$

The final goal is to find the optimal weight parameters α^*, and the specific selection of the optimal question can be determined using the Softmax function: $q_t^m = \sum_{q \in V} \frac{\exp(\alpha_q^{(t,m)})}{\sum_{q' \in V} \exp(\alpha_{q'}^{(t,m)})} q$. Here, q refers to the parameters of the question, such as difficulty and discrimination.

Thus, by directly optimizing this scalable and flexible ATS framework from large-scale data, it avoids the need for manual design and enables the adaptability to complex testing scenarios. When we use IRT as a cognitive diagnosis model, we can leverage the strongly convex properties of IRT (Lord, 2012) to obtain some desirable properties of this optimization problem, as shown below:

Lemma 4.2. *When using IRT as the student response function, the inner function (Eq.6) and the outer function (Eq.5) have the following properties:*

1. *The inner function is strongly convex with respect to θ, and its derivatives with respect to both θ and α are L-Lipschitz continuous. Its second derivatives are also Lipschitz. Specifically $\nabla_\alpha \nabla_\theta L(\theta, D_t(\alpha))$ is τ-Lipschitz continuous and $\nabla_\theta^2 L(\theta, D_t(\alpha))$ is ρ-Lipschitz continuous.*
2. *The outer function is M-Lipschitz continuous, with its derivative with respect to θ being L-Lipschitz continuous, and its derivative with respect to α being W-Lipschitz continuous.*

All proofs of the lemma can be found in Appendix. Lemma 4.2.1 describes the properties of the inner function, which represents the modeling of the student's abilities in the assessment process. The property of strong convexity ensures the monotonicity and convergence of the function within a local range, which is crucial for the convergence speed of the optimization algorithm. Additionally, the Lipschitz continuity with respect to parameters guarantees the smoothness and stability of the function, enabling the optimization algorithm to efficiently search for the optimal solution in parameter space. Lemma 4.2.2 describes the properties of the outer function, which represents the overall optimization goal in the evaluation process. Its Lipschitz continuity ensures the stability and feasibility of the optimization process, while the continuity of the derivative with respect to parameters ensures the effectiveness of the gradient descent algorithm in parameter space.These excellent properties provide guarantees for our optimization theory.

4.2. Approximation and Optimization

The search process of Defination 4.1 considers both the global problem of parameter selection and the abilities of students. It aims to optimize the final outcome, which is established under the condition of known global information. However, directly solving the above problem is infeasible because ATS is sequential, meaning that the response to question $t+1$ depends on the response to question t. In practical applications, future response data cannot be obtained, which means it is impossible to directly solve the entire structure.

To address this issue, we approximate global structure search problem with greedy hierarchical structure search. Global search involves obtaining the optimal solution for each step of the test simultaneously, while Greedy hierarchical structure search aim to find the most suitable problem at the current moment. Specifically, at the t-th step of the test, with $t \in [1, 2, \ldots, T]$, the student's ability estimate is θ_t, and the question selection parameters are α_t.First, update the current student's ability based on the student's previous steps of response until convergence.

$$\theta_t^k = \theta_t^{k-1} - \gamma \nabla_\theta L(\theta, D_t(\alpha_t)) \tag{7}$$

Here k is the iteration identifier. Given the learning rate γ, Eq.(7) will performed until convergence to θ_t^*. This is the optimal solution to obtain student ability estimates at step t. Based on the converged optimal result θ_t^*, the current topic selection parameters are updated. It is worth noting that training the inner function until convergence can incur a significant cost. *Therefore, instead of training until convergence to the optimal solution, we use K times of gradient updates a an approximation.* Using θ_t^K instead of θ_t^*, expenses will be greatly reduced.

The question selection parameters also follow the gradient descent form, as shown in Eq.(8):

$$\alpha_{t+1} = \alpha_t - \beta \nabla_\alpha L(\theta_t^K(\alpha_t), D_u) \tag{8}$$

Here β is the learning rate of α. After the update of the gradient descent, we obtain the new parameter α_t. The next question will be selected based on α_t and added to D_t, enabling the greedy hierarchical search process. Repeating these steps for training is a feasible optimization solution.

However, replacing θ_t^* with θ_t^K will introduces some error (Ji et al., 2021). We provide an upper bound on the error of this approximation algorithm in the following discussion.

Theorem 4.3 (Gradient Estimation Error Bound)**.** *Assume that the student response function is μ-strongly convex (e.g.,IRT). At t step of the test, let B be the upper bound of $\|\theta_t^0(\alpha_t) - \theta_t^*(\alpha_t)\|$. Choose the learning rate $\gamma \leq \frac{1}{L}$. Then we have*

$$\begin{aligned}
&\|\nabla_\alpha L(\theta_t^K(\alpha_t), D_u) - \nabla_\alpha L(\theta_t^*(\alpha_t), D_u)\| \leq \\
&\frac{B}{\mu}\left(L^2(1-\gamma\mu)^K - \frac{M(\tau\mu + L\rho)}{\mu}(1-\gamma\mu)^{\frac{K+1}{2}}\right) \\
&\quad + \frac{ML(1-\gamma\mu)^K}{\mu} + \frac{BM\,(\tau\mu + L\rho)}{\mu^2}(1-\gamma\mu)
\end{aligned} \tag{9}$$

Theorem 4.3 reveals that the estimation error of the gradient decays exponentially with the number of inner loop iterations, denoted as K. The proof of Theorem 4.3 provides a direct description of the convergence rate of the gradient sequence to $\nabla_\alpha L(\theta_t^*(\alpha_t), D_u)$ along the gradient descent path for all corresponding points.And it also demonstrates that as K increases, the error tends to decrease.

Algorithm 1 describes the complete training process of the UATS framework. It combines the dynamic question selection algorithm and the greedy hierarchical structure search mentioned. The goal is to optimize the definition 4.1 At the t step of the test, there is already a sequence D_t. First, select question Q based on the current α_t and obtain the corresponding feedback Y_t. After adding to D_t, update the current student's estimated ability θ_t based on D_t. Then, θ_t is used to update the selection parameters α_{t+1}, and the next question will be selected based on α_{t+1}. This is the entire training process.

Algorithm 1 UATS Framework Training Process

Require: Learning rate γ, β, Initialize parameters θ_0, α_0
```
while not converged do
    Randomly sample a mini-batch of students with query set D_u and initialize support set D_t = {}
    for t = 1, 2, ..., T do
        Choose question Q_t based on α_t : Q_t ~ π(α_t)
        The student's response to Q_t is Y_t
        D_t = D_t ∪ {Q_t, Y_t}
        Set θ_t^0 = θ_{t-1}^K
        for k = 1, 2, ..., K do
            θ_t^k ← θ_t^{k-1} − γ∇_θ L(θ_t^{k-1}, D_t(α_t))
        end for
        α_{t+1} ← α_t − β∇_α L(θ_t^K(α_t), D_u)
    end for
end while
```

In order to further guarantee the properties of the greedy layer-wise search algorithm, we propose the following theory, which ensures that the gradients of the algorithm during the training process converge.

Theorem 4.4 (Convergence Analyze)**.** *Choose $\gamma \leq \frac{1}{L}, \beta = \frac{1}{4W}$. Then the average value of the outer function gradient satisfies the following theorem.*

$$\begin{aligned}
&\frac{1}{T}\sum_{t=0}^{T-1} \|\nabla L(\theta_t^*(\alpha_t), D_u)\|^2 \leq \\
&\frac{16W}{T}(L(\theta_0^*(\alpha_0), D_u) - L(\theta_T^*(\alpha_T), D_u)) + C
\end{aligned} \tag{10}$$

Based on the theorem 4.4, we prove that the average of T gradients of the outer function has an upper bound, which guarantees the property of convergence.

5. Experiments

In this section, we conduct qualitative and quantitative analyses on our UATS framework . This evaluation was carried out on three real-world datasets to assess the effectiveness of our approach.The specific code can be found at: `https://github.com/bigdata-ustc/UATS`. We will provide continuous maintenance.

5.1. Experimental Settings

Datasets. We worked with three distinct sets of educational data: the ASSIST dataset (Pardos et al., 2013), which

Table 1. Statistics of the datasets

Dataset	ASSISTments	NIPS-EDU	EXAM
Students	2.3k	22.0k	9.2K
Questions	26.7k	27.6k	1.6k
Interactions	325k	15M	133K

comprises student practice logs related to mathematics and knowledge concepts problems from the ASSISTments online tutoring system. The NIPS-EDU dataset (Wang et al., 2021) originated from the NeurIPS 2020 Education Challenge and consists of student responses to questions on the NIPS-EDU educational platform. The EXAM dataset was supplied by iFLYTEK Co.,Ltd. which collected the records of junior high school students on mathematical exams. The specific details of these three datasets are given in table 1.

Data Partition and Evaluation Methods. We conducted 5-fold cross-validation on all datasets. For each fold, we allocated 60% of the students for training, 20% for validation, and 20% for testing. In each fold, we employ an early stopping strategy using the validation set to fine-tune the parameters for each method. To mitigate overfitting, we randomly shuffled these partitions at the beginning of each training epoch. The performance metrics for the evaluation included Accuracy (Gao et al., 2021) and the Area Under the Receiver Operating Characteristic Curve (AUC) (Bradley, 1997). All experiments were run on an NVIDIA V100 GPU.

Compared Approaches. When evaluating question selection algorithms in CAT, it is crucial to align them with appropriate cognitive diagnostic models. In our study, we compare the effectiveness of two models: Item Response Theory (IRT) and the Neural Cognitive Diagnostic Model (NeuralCDM) (Wang et al., 2023a). NeuralCDM can cover many cognitive diagnosis models,such as MIRT (Reckase, 2009) and MF (Toscher & Jahrer, 2010; Desmarais, 2012). The following selection algorithms are used for comparison

- **Random**: This method randomly selects questions and serves as a reference for improvement compared to several baselines.
- **FSI** (Lord, 2012): It utilizes maximum Fisher information to select questions. It is effective only when cognitive diagnostic models is IRT
- **KLI** (Chang & Ying, 1996): It utilize Kullback-Leibler information to select questions.It also depends on IRT
- **MAAT** (Bi et al., 2020): It employs an active learning (Krishnakumar, 2007) approach to measure question informativeness to select questions
- **BECAT** (Zhuang et al., 2023): It transforms question selection into the selection of optimal subset.

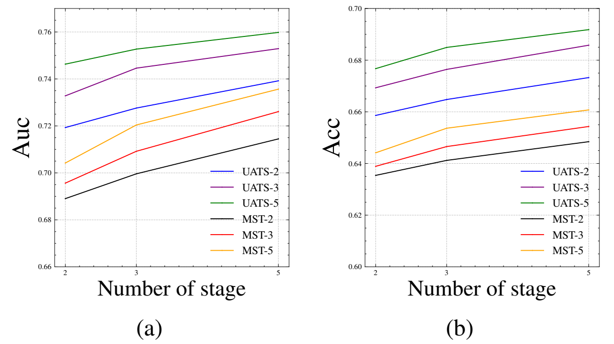


Figure 3. (a) and (b) analyze the AUC and ACC curves during the testing stage for 2, 3, and 5, with a maximum number of 2, 3, and 5 questions per stage. For example, MST-5 means selecting five questions at each stage.

- **NCAT** (Zhuang et al., 2022): It employs a reinforcement learning method to train a DQN architecture for question selection.
- **BOBCAT** (Ghosh & Lan, 2021): It adopts meta learning to learn questions selection algorithms.

5.2. Results and Discussion

In this section, we compare the performance of the two classic adaptive testing tasks to evaluate the effectiveness and efficiency of our proposed UATS framework.

5.2.1. Student Score Prediction

(1) Compare with CAT Baseline. Table 2 presents a comparison of our UATS algorithm with other model designs in the context of adaptive testing. We provide the accuracy (ACC) and area under the curve (AUC) metrics for test lengths of $T = 5$, 10, and 20 as evaluation measures. Our UATS framework achieves the best overall performance on these three datasets. When using the IRT model, the EXAM dataset performs the best. Compared to the original SOTA algorithm, the relative improvement in AUC@20 is on average 1.31%, and for ACC@20, it is 0.71%. This result demonstrates that our UATS framework can provide an accurate ability estimate at the end of the examination.

(2) Compare with MST Baseline. The MST dataset is strictly confidential, and we have no way of accessing it. Therefore we refer to the automatic question generation method of Automated Test Assmbly (ATA) (Breithaupt & Hare, 2007)to create the test modules for MST. Based on the NIPS-EDU dataset, we provide test results for the total number of stages being 2, 3, and 5. The number of modules for each stage is 2, and the number of questions for each module is 2, 3, 5. Figure 3 presents a comparison of UATS with MST using accuracy (ACC) and area under the curve (AUC) as

Table 2. Prediction performance of different methods on ACC and AUC metrics for student achievement prediction. ”-” indicates that information-based selection algorithms cannot be applied to NeuralCDM model. The bold text indicates statistically significant superiority (p-value ≤ 0.01) over the best baseline.

(a) Performances on ASSIST

	IRT			NeuralCDM		
Metric@Step	ACC/AUC@5	ACC/AUC@10	ACC/AUC@20	ACC/AUC@5	ACC/AUC@10	ACC/AUC@20
Random	65.54/66.37	65.72/66.51	66.13/66.80	64.99/66.20	65.18/66.37	65.60/66.66
FSI	65.76/66.51	66.16/66.78	66.74/67.23	–	–	–
KLI	65.76/66.48	66.13/66.76	66.75/67.21	–	–	–
MAAT	65.65/66.46	65.90/66.77	66.18/67.07	65.15/66.12	65.34/66.41	65.64/66.59
BECAT	65.69/66.44	66.12/66.67	66.67/67.19	65.08/66.35	65.54/66.53	65.89/66.85
BOBCAT	65.99/67.45	66.12/67.83	66.63/67.87	66.76/67.03	67.52/67.87	68.20/68.05
NCAT	66.07/67.33	66.23/67.63	66.42/67.88	66.54/66.98	67.21/67.75	68.18/67.32
UATS	**66.13/67.62**	**66.47/67.91**	**67.07/68.00**	**67.16/67.95**	**67.61/68.58**	**68.32/68.92**

(b) Performances on NIPS-EDU

	IRT			NeuralCDM		
Metric@Step	ACC/AUC@5	ACC/AUC@10	ACC/AUC@20	ACC/AUC@5	ACC/AUC@10	ACC/AUC@20
Random	63.17/68.49	64.24/69.88	65.66/72.19	63.57/68.12	64.46/69.83	66.28/72.16
FSI	63.70/69.26	64.74/71.08	66.19/73.48	–	–	–
KLI	63.67/69.22	64.72/71.06	66.26/73.57	–	–	–
MAAT	62.86/68.15	63.01/69.33	63.11/71.39	62.79/67.92	62.97/69.33	64.04/71.02
BECAT	63.40/68.52	64.66/69.90	65.85/72.73	62.81/67.98	64.13/69.56	66.52/73.23
BOBCAT	**64.71**/70.04	**66.07**/71.61	67.51/73.36	63.97/68.80	**69.02**/75.47	69.46/76.02
NCAT	64.37/69.36	65.55/71.33	66.20/72.88	63.27/67.81	66.87/72.76	68.95/76.34
UATS	63.37/**71.79**	65.98/**73.46**	**68.17/74.85**	**64.43/69.79**	68.94/**75.65**	**70.15/76.86**

(c) Performances on EXAM

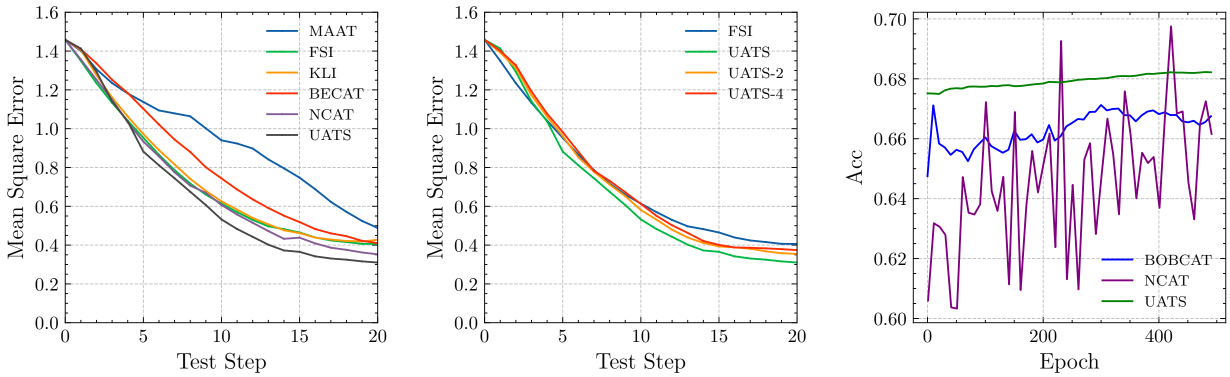
	IRT			NeuralCDM		
Metric@Step	ACC/AUC@5	ACC/AUC@10	ACC/AUC@20	ACC/AUC@5	ACC/AUC@10	ACC/AUC@20
Random	67.72/74.31	68.42/74.92	69.01/75.47	69.53/74.04	70.12/74.92	70.39/75.38
FSI	68.69/75.08	69.52/75.71	70.40/76.75	–	–	–
KLI	68.75/75.07	69.55/75.73	70.63/76.81	–	–	–
MAAT	68.35/74.72	69.79/75.32	69.98/76.24	70.17/74.84	70.51/75.24	71.37/76.26
BECAT	68.49/74.94	69.22/75.59	70.37/76.66	70.15/74.91	70.82/75.52	71.76/76.54
BOBCAT	72.28/76.03	72.57/76.55	73.80/77.32	72.39/76.19	73.62/76.55	74.08/77.58
NCAT	71.89/76.13	72.41/76.67	73.18/77.19	72.18/76.28	73.69/76.73	74.21/77.20
UATS	**72.74/77.81**	**72.96/78.15**	**74.51/78.63**	**73.51/76.78**	**74.12/77.16**	**74.53/77.98**

evaluation metrics. The results show that our UATS framework achieves the best overall performance on these three datasets. It outperforms all other MST methods with an average improvement of 2.7% in ACC and 3.1% in AUC. This result suggests that our UATS framework can perfectly incorporate MST and provide more accurate ability estimation while offering flexibility.

5.2.2. Simulation of Ability Estimation

The simulation of ability estimation is a foundational evaluation technique in ATS (Vie et al., 2017). The goal of adaptive testing is to accurately estimate the abilities of students. We conducted a simulation experiment on the NIPS-EDU dataset to estimate abilities. Specifically, we used the mean squared error $E[\|\theta_t - \theta_0\|^2]$ between the true ability of a student θ_0 and the ability at step t, θ_t. Since the true ability θ_0 is unknown, we approximated it using the feedback from a student answering all the questions in the question bank.

Figure 4(a) shows the metrics of different methods based on the IRT model on the NIPS-EDU dataset for steps 1 to 20. As the number of selected questions increases, we find that the UATS method consistently achieves a lower estimation error. Some implicit methods that do not prioritize estimation accuracy (e.g., NCAT) perform better in the initial stages, but still lag behind the UATS framework after 4 step. Compared to the SOTA algorithm, UATS can achieve the same estimation error with fewer questions. On aver-



(a) (b) (c)

Figure 4. (a) demonstrates the comparison between different baselines and our method UATS on the NIPS dataset. (b) shows the impact of different numbers of question pools in one layer on student ability estimation.(c) illustrates the comparison between UATS framework and other data-driven methods for the first 500 epochs.

age, it can achieve the same estimation accuracy with 20% fewer questions, demonstrating its efficiency in estimating abilities, that is, reducing the length of the test.

We also found that the number of question packs in the UATS framework affects the estimation of ability. It can be seen from the Figure 4(b) that as the number of question packs increases, the same step mean squared error becomes higher. This is because more question packs imply a closer approach to MST, which requires more questions to be answered. When the number of question packs is small, it becomes closer to CAT, which means fewer questions are answered in one stage and the adaptability is stronger. Therefore, effective control of the number of question packs can achieve a balance between MST and CAT.

5.2.3. Convergence property analysis

In order to evaluate the convergence of the proposed method, we compared the training curves with several popular data-driven methods. Figure 4(c) shows the training curves for the first 500 epochs the ASSISTment dataset. It can be observed that our method not only achieves the best results but also steadily improves during training. Compared to other methods, it reduces fluctuations and achieves a more stable convergence. The results indicate that our method demonstrates smoother convergence and higher stability while also exhibiting relatively controllable variations in performance metrics during training.

6. Conclusions

In this paper, we proposed a UATS framework, which provided a unified method for learning selection algorithms from real-world data in online education. This framework covered existing adaptive testing frameworks, including computerized adaptive testing and multistage testing. Specifically, in order to overcome the limitations of CAT on students' test-taking and MST on expert preset requirements, we modeled them as a structure search problem and solved the limitations in an effective deep learning environment. In addition, we proposed a dynamic and differentiable selection algorithm. We transformed global search into a greedy search problem and provided a proof of convergence. Extensive experiments showed that the UATS framework was compatible with adaptability and flexibility. It accurately measured students' proficiency and reduced the length of the test. Moreover, our method exhibited greater stability during training compared to other data-driven frameworks.

Impact Statement

In adaptive testing systems, different students may be recommended different questions, raising concerns about fairness. Our paper focuses on proposing a novel adaptive testing method, while fairness is another independent research area (Lord, 2012; Zhang et al., 2024), so it is beyond the scope of our discussion.

Acknowledgment

This research was partial supported by grants from the National Key Research and Development Program of China (Grant No.2021YFF0901003) and the National Natural Science Foundation of China (Grants No.62106244, No.U20A20229, and No.62172320), the University Synergy Innovation Program of Anhui Province (No.GXXT-2022-042)

References

Ackerman, T. A., Gierl, M. J., and Walker, C. M. Using multidimensional item response theory to evaluate educational and psychological tests. *Educational Measurement: Issues and Practice*, 22(3):37–51, 2003.

Bi, H., Ma, H., Huang, Z., Yin, Y., Liu, Q., Chen, E., Su, Y., and Wang, S. Quality meets diversity: A model-agnostic framework for computerized adaptive testing. In *2020 IEEE International Conference on Data Mining (ICDM)*, pp. 42–51. IEEE, 2020.

Bradley, A. P. The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern recognition*, 30(7):1145–1159, 1997.

Breithaupt, K. and Hare, D. R. Automated simultaneous assembly of multistage testlets for a high-stakes licensing examination. *Educational and Psychological Measurement*, 67(1):5–20, 2007.

Chang, H.-H. Psychometrics behind computerized adaptive testing. *Psychometrika*, 80(1):1–20, 2015.

Chang, H.-H. and Ying, Z. A global information approach to computerized adaptive testing. *Applied Psychological Measurement*, 20(3):213–229, 1996.

Chen, S. J., Choi, A., and Darwiche, A. Computer adaptive testing using the same-decision probability. In *BMA@ UAI*, pp. 34–43, 2015.

Cheng, S., Liu, Q., Chen, E., Huang, Z., Huang, Z., Chen, Y., Ma, H., and Hu, G. Dirt: Deep learning enhanced item response theory for cognitive diagnosis. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pp. 2397–2400, 2019.

Desmarais, M. C. Mapping question items to skills with non-negative matrix factorization. *ACM Sigkdd Explorations Newsletter*, 13(2):30–36, 2012.

Embretson, S. E. and Reise, S. P. *Item response theory*. Psychology Press, 2013.

Feng, W., Ghosh, A., Sireci, S., and Lan, A. S. Balancing test accuracy and security in computerized adaptive testing. In *International Conference on Artificial Intelligence in Education*, pp. 708–713. Springer, 2023.

Gao, L., Zhao, Z., Li, C., Zhao, J., and Zeng, Q. Deep cognitive diagnosis model for predicting students' performance. *Future Generation Computer Systems*, 126: 252–262, 2022.

Gao, W., Liu, Q., Huang, Z., Yin, Y., Bi, H., Wang, M.-C., Ma, J., Wang, S., and Su, Y. Rcd: Relation map driven cognitive diagnosis for intelligent education systems. In *Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval*, pp. 501–510, 2021.

Ghosh, A. and Lan, A. Bobcat: Bilevel optimization-based computerized adaptive testing. In Zhou, Z.-H. (ed.), *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, pp. 2410–2417. International Joint Conferences on Artificial Intelligence Organization, 8 2021. doi: 10.24963/ijcai.2021/332. URL https://doi.org/10.24963/ijcai.2021/332. Main Track.

Gilavert, P. and Freire, V. Computerized adaptive testing: A unified approach under markov decision process. In *International Conference on Computational Science and Its Applications*, pp. 591–602. Springer, 2022.

Huang, Z., Liu, Q., Chen, Y., Wu, L., Xiao, K., Chen, E., Ma, H., and Hu, G. Learning or forgetting? a dynamic approach for tracking the knowledge proficiency of students. *ACM Transactions on Information Systems (TOIS)*, 38(2):1–33, 2020.

Ji, K., Yang, J., and Liang, Y. Provably faster algorithms for bilevel optimization and applications to meta-learning, 2021. URL https://openreview.net/forum?id=sMEpviTLi1h.

Kang, H.-A., Zhang, S., and Chang, H.-H. Dual-objective item selection criteria in cognitive diagnostic computerized adaptive testing. *Journal of Educational Measurement*, 54(2):165–183, 2017.

Krishnakumar, A. Active learning literature survey. *Tech. rep., Technical reports, University of California, Santa Cruz.*, 42, 2007.

Li, X., Xu, H., Zhang, J., and Chang, H.-h. Deep reinforcement learning for adaptive learning systems. *Journal of Educational and Behavioral Statistics*, 48(2):220–243, 2023.

Liu, Q., Zhuang, Y., Bi, H., Huang, Z., Huang, W., Li, J., Yu, J., Liu, Z., Hu, Z., Hong, Y., et al. Survey of computerized adaptive testing: A machine learning perspective. *arXiv preprint arXiv:2404.00712*, 2024.

Lord, F. M. *Applications of item response theory to practical testing problems*. Routledge, 2012.

Ma, H., Zeng, Y., Yang, S., Qin, C., Zhang, X., and Zhang, L. A novel computerized adaptive testing framework with decoupled learning selector. *Complex & Intelligent Systems*, 9(5):5555–5566, 2023.

Magis, D., Yan, D., and Von Davier, A. A. *Computerized adaptive and multistage testing with R: Using packages catR and mstR*. Springer, 2017.

Mead, A. D. An introduction to multistage testing. *Applied Measurement in Education*, 19(3):185–187, 2006.

Meijer, R. R. and Nering, M. L. Computerized adaptive testing: Overview and introduction. *Applied psychological measurement*, 23(3):187–194, 1999.

Mujtaba, D. F. and Mahapatra, N. R. Multi-objective optimization of item selection in computerized adaptive testing. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 1018–1026, 2021.

Nurakhmetov, D. Reinforcement learning applied to adaptive classification testing. *Theoretical and practical advances in computer-based educational measurement*, pp. 325–336, 2019.

Pardos, Z. A., Baker, R. S., San Pedro, M. O., Gowda, S. M., and Gowda, S. M. Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes. In *Proceedings of the third international conference on learning analytics and knowledge*, pp. 117–124, 2013.

Reckase, M. D. *Multidimensional Item Response Theory*. Springer, 2009.

Rudner, L. M. An examination of decision-theory adaptive testing procedures. In *annual meeting of the American Educational Research Association*, 2002.

Sari, H. I., Yahsi-Sari, H., and HUGGINS-MANLEY, A. C. Computer adaptive multistage testing: Practical issues, challenges and principles. *Journal of Measurement and Evaluation in Education and Psychology*, 7(2):388–406, 2016.

Shen, J., Qian, H., Zhang, W., and Zhou, A. Symbolic cognitive diagnosis via hybrid optimization for intelligent education systems. In *Proceedings of the 38th AAAI Conference on Artificial Intelligence*, pp. 14928–14936, Vancouver, Canada, 2024.

Toscher, A. and Jahrer, M. Collaborative filtering applied to educational data mining. *KDD cup*, 2010.

van der Linden, W. J. Bayesian item selection criteria for adaptive testing. *Psychometrika*, 63(2):201–216, 1998.

Van der Linden, W. J. and Pashley, P. J. Item selection and ability estimation in adaptive testing. In *Elements of adaptive testing*, pp. 3–30. Springer, 2009.

Veerkamp, W. J. and Berger, M. P. Some new item selection criteria for adaptive testing. *Journal of Educational and Behavioral Statistics*, 22(2):203–226, 1997.

Veldkamp, B. P. and Verschoor, A. J. Robust computerized adaptive testing. *Theoretical and practical advances in computer-based educational measurement*, pp. 291–305, 2019.

Vie, J.-J., Popineau, F., Bruillard, É., and Bourda, Y. A review of recent advances in adaptive assessment. *Learning analytics: fundaments, applications, and trends*, pp. 113–142, 2017.

Wang, F., Liu, Q., Chen, E., Huang, Z., Yin, Y., Wang, S., and Su, Y. Neuralcd: A general framework for cognitive diagnosis. *IEEE Transactions on Knowledge and Data Engineering*, 35(8):8312–8327, 2023a. doi: 10.1109/TKDE.2022.3201037.

Wang, H., Long, T., Yin, L., Zhang, W., Xia, W., Hong, Q., Xia, D., Tang, R., and Yu, Y. Gmocat: A graph-enhanced multi-objective method for computerized adaptive testing. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 2279–2289, 2023b.

Wang, Z., Lamb, A., Saveliev, E., Cameron, P., Zaykov, J., Hernandez-Lobato, J. M., Turner, R. E., Baraniuk, R. G., Craig Barton, E., Peyton Jones, S., Woodhead, S., and Zhang, C. Results and insights from diagnostic questions: The neurips 2020 education challenge. In Escalante, H. J. and Hofmann, K. (eds.), *Proceedings of the NeurIPS 2020 Competition and Demonstration Track*, volume 133 of *Proceedings of Machine Learning Research*, pp. 191–205. PMLR, 06–12 Dec 2021. URL https://proceedings.mlr.press/v133/wang21a.html.

Yu, J., Zhenyu, M., Lei, J., Yin, L., Xia, W., Yu, Y., and Long, T. Sacat: Student-adaptive computerized adaptive testing. In *Proceedings of the Fifth International Conference on Distributed Artificial Intelligence*, pp. 1–7, 2023.

Zhang, Z., Wu, L., Liu, Q., Liu, J., Huang, Z., Yin, Y., Zhuang, Y., Gao, W., and Chen, E. Understanding and improving fairness in cognitive diagnosis. *Science China Information Sciences*, 67(5):152106, 2024.

Zhuang, Y., Liu, Q., Huang, Z., Li, Z., Shen, S., and Ma, H. Fully adaptive framework: Neural computerized adaptive testing for online education. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(4):4734–4742, Jun. 2022.

Zhuang, Y., Liu, Q., Zhao, G., Huang, Z., Huang, W., Pardos, Z., Chen, E., Wu, J., and Li, X. A bounded ability estimation for computerized adaptive testing. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.

A. Proofs of Lemma 4.2

Lemma 4.2 *When using IRT as the student response function, the inner function (Eq.6) and the outer function (Eq.5) have the following properties:*

1. *The inner function is strongly convex with respect to θ, and its derivatives with respect to both θ and α are L-Lipschitz continuous. Its second derivatives are also Lipschitz. Specifically $\nabla_\alpha \nabla_\theta L(\theta, D_t(\alpha))$ is τ-Lipschitz and $\nabla_\theta^2 L(\theta, D_t(\alpha))$ is ρ-Lipschitz continuous.*

2. *The outer function is M-Lipschitz continuous, with its derivative with respect to θ being L-Lipschitz continuous, and its derivative with respect to α being W-Lipschitz continuous.*

Proofs of Lemma 4.2.1

Proof. First,we prove inner function the strongly convex w.r.t θ. We expand the inner function and substitute the loss function with Definition 4.1, yielding

$$L(\theta, D_t(\alpha)) = -\sum_{i=1}^{t}(y_i log(\sigma(\theta - b_i) + (1 - y_i)log(1 - \sigma(\theta - b_i)) \tag{11}$$

Since we use IRT as the response function, which can be defined as: $\sigma(\theta - b_i)$. It is easy to know that IRT is a strongly convex function. Since the cross-entropy loss is strongly convex ,when we calculate the loss using the IRT model and sum them up, the inner function w.r.t θ is also guaranteed to be strongly convex.

Secondly,we prove the inner function's derivatives with respect to both θ and α are L-Lipschitz continuous.

$$\begin{aligned}\nabla_\theta L(\theta, D_t(\alpha)) &= -\sum_{i=1}^{t} \nabla_\theta(y_i \log(\sigma(\theta - b_i)) + (1 - y_i)\log(1 - \sigma(\theta - b_i)) \\ &= -\sum_{i=1}^{S} y_i - \sigma(\theta(\alpha) - b_i)\end{aligned} \tag{12}$$

$$\nabla_\alpha L(\theta, D_t(\alpha)) = -\left(\sum_{i=1}^{t-1} \frac{\partial\theta(\alpha)}{\partial\alpha}(y_i - \sigma(\theta(\alpha) - b_i)) + (\frac{\partial\theta(\alpha)}{\partial\alpha} - \frac{\partial b_t}{\partial\alpha})(y_t - \sigma(\theta(\alpha) - b_t))\right) \tag{13}$$

The derivative of θ can be easily seen as bounded, therefore it is L-Lipschitz continuous. The derivative of α exists as a variable. $\frac{\partial\theta(\alpha)}{\partial\alpha}$ and $\frac{\partial b_t}{\partial\alpha}$. $\frac{\partial\theta(\alpha)}{\partial\alpha}$ is the gradient for updating the inner function and always has a clear upper bound. $\frac{\partial b_t}{\partial\alpha}$ is the gradient for the selected topic by the dynamic topic selection algorithm with respect to the topic selection parameter, and it is also guaranteed to be bounded. Therefore it is L-Lipschitz

Finally ,we prove the inner function's second derivatives are also Lipschitz. Eq.(12) represents the first-order derivative of the inner function with respect to θ. Based on this, we have the following proof:

$$\begin{aligned}\nabla_\alpha\nabla_\theta L(\theta, D_t(\alpha)) &= -\sum_{i=1}^{t-1}(\nabla_\alpha(y_i - \sigma(\theta(\alpha) - b_i)) + \nabla_\alpha(y_t - \sigma(\theta(\alpha) - b_t))) \\ &= \sum_{i=1}^{t-1}\left(\frac{\partial\theta(\alpha)}{\partial\alpha}(\sigma(\theta(\alpha) - b_i))(1 - \sigma(\theta(\alpha) - b_i)) + (\frac{\partial\theta(\alpha)}{\partial\alpha} - \frac{\partial b_t}{\partial\alpha})(\sigma(\theta(\alpha) - b_t))(1 - \sigma(\theta(\alpha) - b_t))\right)\end{aligned} \tag{14}$$

$\frac{\partial\theta}{\partial\alpha}$ and $\frac{\partial b_t}{\partial\alpha}$ is guaranteed to be bounded. Therefore, it can be proven that the overall function is bounded, satisfying τ-Lipschitz

Next, we will prove $\nabla^2_\theta L(\theta, D_t(\alpha))$ is ρ-Lipschitz continuous:

$$\nabla^2_\theta L(\theta, D_t(\alpha)) = -\sum_{i=1}^{t} \nabla_\theta (y_i - \sigma(\theta(\alpha) - b_i)) \tag{15}$$

$$= \sum_{i=1}^{t} (1 - \sigma(\theta(\alpha) - b_i))(\sigma(\theta(\alpha) - b_i)) \tag{16}$$

It is obvious that this is bounded. We complete the proof. □

Proofs of Lemma 4.2.2

Proof. First, we prove the outer function is M-Lipschitz. Combined with Definition 4.1, represents the loss sum of a student's answers to questions in the query set, denoted as $L(\theta(\alpha), D_u)$.

$$L(\theta(\alpha), D_u) = -\sum_{i=1}^{U} (y_i log(\sigma(\theta - b_i) + (1 - y_i) log(1 - \sigma(\theta - b_i)) \tag{17}$$

It is evident that this is bounded, which indicates that we can always find an M that satisfies Lipschitz.

Next, we prove the derivative of the outer function $L(\theta(\alpha), D_u)$ with respect to θ is L-Lipschitz:

$$\nabla_\theta L(\theta(\alpha), D_u) = -\sum_{i=1}^{U} \nabla_\theta (y_i \log(\sigma(\theta(\alpha) - b_i)) + (1 - y_i) \log(1 - \sigma(\theta(\alpha) - b_i))$$

$$= -\sum_{i=1}^{U} y_i(1 - \sigma(\theta(\alpha) - b_i)) - (1 - y_i)(\sigma(\theta(\alpha) - b_i))$$

$$= -\sum_{i=1}^{U} y_i - \sigma(\theta(\alpha) - b_i) \tag{18}$$

And the Eq.(18) is definitely bounded, so L-Lipschitz is certainly satisfied.

Finaly we prove the derivative of the outer function $L(\theta(\alpha), D_u)$ with respect to θ is W-Lipschitz: The expansion of the derivative of the loss function is as follows:

$$\nabla_\alpha L(\theta^*(\alpha), D_u) = \sum_{i=1}^{U} \nabla_\alpha (y_i \log(\sigma(\theta^*(\alpha) - b_i)) + (1 - y_i) \log(1 - \sigma(\theta^*(\alpha) - b_i))$$

$$= \frac{\partial \theta^*(\alpha)}{\partial \alpha} \sum_{i=1}^{U} y_i(1 - \sigma(\theta^*(\alpha) - b_i)) - (1 - y_i)(\sigma(\theta^*(\alpha) - b_i)$$

$$= \frac{\partial \theta^*(\alpha)}{\partial \alpha} \sum_{i=1}^{U} y_i - \sigma(\theta^*(\alpha) - b_i) \tag{19}$$

Eq.(19) is bounded. It can be proven that the derivative with respect to α_t is bounded. Thus, the proof is complete. □

B. Proofs of Theorem 4.3

Theorem 4.3 (Gradient Estimation Error Bound). *Assume that the student response function is μ-strongly convex (e.g.IRT). At t step of the test, let D be the upper bound of $\|\theta_t^0(\alpha_t) - \theta_t^*(\alpha_t)\|$ .Choose the learning rate $\gamma \leq \frac{1}{L}$.we have*

$$\begin{aligned} &\|\nabla_\alpha L(\theta_t^K(\alpha_t), D_u) - \nabla_\alpha L(\theta_t^*(\alpha_t), D_u)\| \leq \\ &\frac{B}{\mu}\left(L^2(1-\gamma\mu)^K - \frac{M(\tau\mu + L\rho)}{\mu}(1-\gamma\mu)^{\frac{K+1}{2}}\right) \\ &\quad + \frac{ML(1-\gamma\mu)^K}{\mu} + \frac{BM\,(\tau\mu + L\rho)}{\mu^2}(1-\gamma\mu) \end{aligned} \tag{20}$$

Proof. Using

$$\nabla_\alpha L(\theta_t^*(\alpha), D_u) = \frac{\partial \theta_t^*(\alpha_t)}{\partial \alpha_t} \nabla_\theta L(\theta_t^*(\alpha_t), D_u) \tag{21}$$

$$\nabla_\alpha L(\theta_t^K(\alpha_t), D_u) = \frac{\partial \theta_t^K(\alpha_t)}{\partial \alpha_t} \nabla_\theta L(\theta_t^K(\alpha_t), D_u) \tag{22}$$

Their difference dis

$$\begin{aligned} &\|\nabla_\alpha L(\theta_t^K(\alpha_t), D_u) - \nabla_\alpha L(\theta_t^*(\alpha_t), D_u)\| \\ &= \| \frac{\partial \theta_t^K(\alpha_t)}{\partial \alpha_t} \nabla_\theta L(\theta_t^K(\alpha_t), D_u) - \frac{\partial \theta_t^*(\alpha_t)}{\partial \alpha_t} \nabla_\theta L(\theta_t^K(\alpha_t), D_u) \\ &+ \frac{\partial \theta_t^*(\alpha_t)}{\partial \alpha_t} \nabla_\theta L(\theta_t^K(\alpha_t), D_u) - \frac{\partial \theta_t^*(\alpha_t)}{\partial \alpha_t} \nabla_\theta L(\theta_t^*(\alpha_t), D_u)\| \end{aligned} \tag{23}$$

Using the triangle inequality

$$\begin{aligned} &\|\nabla_\alpha L(\theta_t^K(\alpha_t), D_u) - \nabla_\alpha L(\theta_t^*(\alpha_t), D_u)\| \\ &\leq \| \frac{\partial \theta_t^K(\alpha_t)}{\partial \alpha_t} - \frac{\partial \theta_t^*(\alpha_t)}{\partial \alpha_t} \| \|\nabla_\theta L(\theta_t^K(\alpha_t), D_u)\| + \| \frac{\partial \theta_t^*(\alpha_t)}{\partial \alpha_t} \| \|\nabla_\theta L(\theta_t^K(\alpha_t), D_u) - \nabla_\theta L(\theta_t^*(\alpha_t), D_u)\| \end{aligned} \tag{24}$$

According to lemma 4.2.2 we have

$$\leq M \left\| \frac{\partial \theta_t^K(\alpha_t)}{\partial \alpha_t} - \frac{\partial \theta_t^*(\alpha_t)}{\partial \alpha_t} \right\| + L \left\| \frac{\partial \theta_t^*(\alpha_t)}{\partial \alpha_t} \right\| \|\theta_t^K(\alpha_t) - \theta_t^*(\alpha_t)\|, \tag{25}$$

Next we will get the upper bound of $\|\frac{\partial \theta_t^K(\alpha_t)}{\partial \alpha_t} - \frac{\partial \theta^*(\alpha_t)}{\partial \alpha_t}\|$
According to the gradient descent.$\theta_t^k = \theta_t^{k-1} - \gamma \nabla_\theta L(\theta_t^{k-1}, D_t(\alpha_t))$ using the chainrule,we have

$$\frac{\partial \theta_t^k(\alpha_t)}{\partial \alpha_t} = \frac{\partial \theta_t^{k-1}(\alpha_t)}{\partial \alpha_t} - \gamma \left(\nabla_\alpha \nabla_\theta L(\theta_t^{k-1}, D_t(\alpha_t)) + \frac{\partial \theta_t^{k-1}}{\partial \alpha_t} \nabla_\theta^2 L(\theta_t^{k-1}, D_t(\alpha_t)) \right) \tag{26}$$

Based on the optimality of $\theta^*(\alpha_t)$,we have$L(\theta_t^*, D_t(\alpha_t))$ So we have

$$\nabla_\alpha \nabla_\theta L(\theta_t^*, D_t(\alpha_t)) + \frac{\partial \theta_t^*(\alpha_t)}{\partial \alpha_t} \nabla_\theta^2 L(\theta_t^*, D_t(\alpha_t)) = 0 \tag{27}$$

Substituting Eq.(27) into Eq.(26) yields

$$\begin{aligned} &\frac{\partial \theta_t^k(\alpha_t)}{\partial \alpha_t} - \frac{\partial \theta_t^*(\alpha_t)}{\partial \alpha_t} \\ &= \frac{\partial \theta_t^{k-1}(\alpha_t)}{\partial \alpha_t} - \frac{\partial \theta_t^*(\alpha_t)}{\partial \alpha_t} - \gamma \left(\nabla_\alpha \nabla_\theta L(\theta_t^{k-1}, D_t(\alpha_t)) + \frac{\partial \theta_t^{k-1}}{\partial \alpha_t} \nabla_\theta^2 L(\theta_t^{k-1}, D_t(\alpha_t)) \right) \\ &\quad + \gamma \left(\nabla_\alpha \nabla_\theta L(\theta_t^*, D_t(\alpha_t)) + \frac{\partial \theta_t^*}{\partial \alpha_t} \nabla_\theta^2 L(\theta_t^*, D_t(\alpha_t)) \right) \\ &= \frac{\partial \theta_t^{k-1}(\alpha_t)}{\partial \alpha_t} - \frac{\partial \theta_t^*(\alpha_t)}{\partial \alpha_t} - \gamma \left(\nabla_\alpha \nabla_\theta L(\theta_t^{k-1}(\alpha_t), D_t(\alpha_t)) - \nabla_\alpha \nabla_\theta L(\theta_t^*, D_t(\alpha_t)) \right) \\ &\quad - \gamma \left(\frac{\partial \theta_t^{k-1}}{\partial \alpha_t} - \frac{\partial \theta_t^*(\alpha_t)}{\partial \alpha_t} \right) \nabla_\theta^2 L(\theta_t^{k-1}, D_t(\alpha_t)) + \gamma \frac{\partial \theta_t^*(\alpha_t)}{\partial \alpha_t} \left(\nabla_\theta^2 L(\theta_t^*, D_t(\alpha_t)) - \nabla_\theta^2 L(\theta_t^{k-1(\alpha_t)}, D_t(\alpha_t)) \right) \end{aligned} \tag{28}$$

Combining Eq.(27) and Lemma 4.2.1 yields

$$\left\| \frac{\partial \theta_t^*(\alpha_t)}{\partial \alpha_t} \right\| = \left\| \nabla_\alpha \nabla_\theta L(\theta_t^*, D_t(\alpha_t)) \left[\nabla_\theta^2 L(\theta_t^*, D_t(\alpha_t)) \right]^{-1} \right\| \leq \frac{L}{\mu}. \tag{29}$$

Then we have

$$\begin{aligned}
&\left\|\frac{\partial\theta_t^k(\alpha_t)}{\partial\alpha_t}-\frac{\partial\theta^*(\alpha_t)}{\partial\alpha_t}\right\| \\
&\le \left\|I-\gamma\nabla_\theta^2 L(\theta_t^{k-1},D_t(\alpha_t))\right\|\left\|\frac{\partial\theta_t^{k-1}(\alpha_t)}{\partial\alpha_t}-\frac{\partial\theta_t^*(\alpha_t)}{\partial\alpha_t}\right\|+\gamma\left(\tau+\frac{L\rho}{\mu}\right)\|\theta_t^{k-1}-\theta^*(\alpha_t)\| \\
&\le (1-\gamma\mu)\|\frac{\partial\theta_t^{k-1}}{\partial\alpha_t}-\frac{\partial\theta^*(\alpha_t)}{\partial\alpha_t}\|+\gamma\left(\tau+\frac{L\rho}{\mu}\right)\|\theta_t^{k-1}-\theta^*(\alpha_t)\|
\end{aligned} \tag{30}$$

Based on the lemma 4.2.1,the strongly-convex lower-level function $L(\theta_t, D_t(\alpha_t))$,we have

$$\|\theta_t^{k-1}(\alpha_t)-\theta^*(\alpha_t)\|\le(1-\gamma\mu)^{\frac{k-1}{2}}\|\theta_t^0(\alpha_t)-\theta^*(\alpha_t)\|. \tag{31}$$

Over k from 1 to K,we have

$$\begin{aligned}
&\left\|\frac{\partial\theta_t^K(\alpha_t)}{\partial\alpha_t}-\frac{\partial\theta_t^*(\alpha_t)}{\partial\alpha_t}\right\| \\
&\le (1-\gamma\mu)^K\left\|\frac{\partial\theta_t^0(\alpha_t)}{\partial\alpha_t}-\frac{\partial\theta_t^*(\alpha_t)}{\partial\alpha_t}\right\|+\gamma\left(\tau+\frac{L\rho}{\mu}\right)\sum_{k=0}^{K-1}(1-\gamma\mu)^{K-1-k}(1-\gamma\mu)^{\frac{k}{2}}\|\theta_t^0(\alpha_t)-\theta_t^*(\alpha_t)\| \\
&= (1-\gamma\mu)^K\left\|\frac{\partial\theta_t^0(\alpha_t)}{\partial\alpha_t}-\frac{\partial\theta_t^*(\alpha_t)}{\partial\alpha_t}\right\|+\frac{(\tau\mu+L\rho)}{\mu^2}(1-\gamma\mu-(1-\gamma\mu)^{\frac{K+1}{2}})\|\theta_t^0(\alpha_t)-\theta_t^*(\alpha_t)\| \\
&\le \frac{L(1-\gamma\mu)^K}{\mu}+\left(-\frac{(\tau\mu+L\rho)}{\mu^2}(1-\gamma\mu)^{\frac{K+1}{2}}+\frac{(\tau\mu+L\rho)}{\mu^2}(1-\gamma\mu)\right)\|\theta_t^0(\alpha_t)-\theta_t^*(\alpha_t)\|
\end{aligned} \tag{32}$$

We have the upper bound of $\|\frac{\partial\theta_t^K(\alpha_t)}{\partial\alpha_t}-\frac{\partial\theta^*(\alpha_t)}{\partial\alpha_t}\|$ .Let B be the upper bound of $\|\theta_t^0(\alpha_t)-\theta_t^*(\alpha_t)\|$The final result is yielded:

$$\begin{aligned}
&\|\nabla_\alpha L(\theta_t^K(\alpha_t),D_u)-\nabla_\alpha L(\theta_t^*(\alpha_t),D_u)\|\le \\
&\frac{L^2}{\mu}(1-\gamma\mu)^{\frac{K}{2}}\|\theta_t^0-\theta^*(\alpha_t)\| \\
&+M\left(\frac{L(1-\gamma\mu)^K}{\mu}+\left(-\frac{(\tau\mu+L\rho)}{\mu^2}(1-\gamma\mu)^{\frac{K+1}{2}}+\frac{(\tau\mu+L\rho)}{\mu^2}(1-\gamma\mu)\right)\|\theta_t^0-\theta^*(\alpha_t)\|\right) \\
&=\frac{B}{\mu}\left(L^2(1-\gamma\mu)^{\frac{K}{2}}-\frac{M(\tau\mu+L\rho)}{\mu}(1-\gamma\mu)^{\frac{K+1}{2}}\right)+\frac{ML(1-\gamma\mu)^K}{\mu}+\frac{BM\,(\tau\mu+L\rho)}{\mu^2}(1-\gamma\mu)
\end{aligned} \tag{33}$$

The prove is complete. □

C. Proofs of Theorem 4.4

Theorem 4.4 (Convergence Analyze). *Choose* $\gamma\le\frac{1}{L},\beta=\frac{1}{4W}$. *Then the average value of the outer function gradient satisfies the following theorem.*

$$\begin{aligned}
&\frac{1}{T}\sum_{t=0}^{T-1}\|\nabla L(\theta_t^*(\alpha_t),D_u)\|^2\le \\
&\frac{16W}{T}(L(\theta_0^*(\alpha_0),D_u)-L(\theta_T^*(\alpha_T),D_u))+C
\end{aligned} \tag{34}$$

Proof. According to lemma 4.2,2,the outer function's derivate w.r.t α is W-Lipschitz. We have

$$\begin{aligned} &L(\theta^*_{t+1}(\alpha_{t+1}), D_u) \leq L(\theta^*_t(\alpha_t), D_u) + \langle \nabla_\alpha L(\theta^*_t(\alpha_t), D_u), \alpha_{t+1} - \alpha_t \rangle + \frac{W}{2}\|\alpha_{t+1} - \alpha_t\|^2 \\ &\leq L(\theta^*_t(\alpha_t), D_u) - \beta\langle \nabla_\alpha L(\theta^K_t(\alpha_t), D_u) - L(\theta^*_t(\alpha_t), D_u)\rangle - \beta\|L(\theta^*_t(\alpha_t), D_u)\|^2 + \beta W\|L(\theta^*_t(\alpha_t), D_u)\|^2 \\ &\quad + \beta^2 W\|\nabla L(\theta^*_t(\alpha_t), D_u) - L(\theta^K_t(\alpha_t), D_u)\|^2 \\ &\leq L(\theta^*_t(\alpha_t), D_u) - \left(\frac{\beta}{2} - \beta^2 W\right)\|\nabla L(\theta^*_t(\alpha_t), D_u)\|^2 + (\frac{\beta}{2} + \beta^2 W)\|\nabla L(\theta^*_t(\alpha_t), D_u) - L(\theta^K_t(\alpha_t), D_u)\|^2 \end{aligned} \tag{35}$$

Substitute theorem 4.3 into the Eq.(35)

$$\begin{aligned} &L(\theta^*_{t+1}(\alpha_{t+1}), D_u) \leq L(\theta^*_t(\alpha_t), D_u) - \left(\frac{\beta}{2} - \beta^2 W\right)\|\nabla L(\theta^*_t(\alpha_t), D_u)\|^2 \\ &+ (\frac{\beta}{2} + \beta^2 W)\frac{B^2}{\mu^2}\left(L^4(1-\gamma\mu)^K. + \frac{M^2(\tau\mu + L\rho)^2}{\mu^2}(1-\gamma\mu - (1-\gamma\mu)^{\frac{K+1}{2}})^2 + \frac{M^2L^2(1-\gamma\mu)^{2K}}{\mu^2}\right) \end{aligned} \tag{36}$$

Telescoping Eq.(36) over t from 0 to T - 1 yields

$$\begin{aligned} &\frac{1}{T}\sum_{t=0}^{T-1}\left(\frac{1}{2} - \beta W\right)\|\nabla L(\theta^*_t(\alpha_t), D_u)\|^2 \\ &\leq \frac{L(\theta^*_0(\alpha_0), D_u) - L(\theta^*_T(\alpha_T), D_u)}{\beta T} + \left(\frac{1}{2} + \beta W\right)\left(\frac{B^2}{\mu^2}(L^4(1-\gamma\mu)^{2K} + \frac{M^2(\tau\mu + L\rho)^2}{\mu^2}(1-\gamma\mu)^{K+1})\right. \\ &\left. + \frac{M^2L^2(1-\gamma\mu)^{2K}}{\mu^2} + \frac{B^2M^2\left(\tau\mu + L\rho\right)^2}{\mu^4}(1-\gamma\mu)^2\right) \end{aligned} \tag{37}$$

Substitute $\beta = \frac{1}{4W}$ and $K = \log\left(\max\left\{\frac{BL^2}{\mu}, \frac{(1-\gamma\mu)M^2(\tau\mu+L\rho)^2B^2}{\mu^4}, \frac{ML}{\mu}\right\}\frac{9}{4\epsilon}\right) / \log\frac{1}{1-\gamma\mu}, C = \frac{3B^2M^2(\tau\mu+L\rho)^2(1-\gamma\mu)^2}{4\mu^4}$ $+\epsilon$ into this Eq.(37). We have

$$\frac{1}{T}\sum_{t=0}^{T-1}\|\nabla L(\theta^*_t(x), D_u)\|^2 \leq \frac{16W}{T}(L(\theta^*_0(\alpha_0), D_u) - L(\theta^*_T(\alpha_T), D_u)) + C \tag{38}$$

The prove is complete. □