BiO-HMC: Dynamic Human-Machine Collaboration for Consensus Decision-Making via Bilevel Optimization

Yinghui Pan[1], Shuaijie Zhao[1], Shenbao Yu[2*], Zongyang Liu[3], Yifeng Zeng[4], Han Liu[5,6], Mingwei Lin[2]

[1] School of Artificial Intelligence & National Engineering Laboratory for Big Data System Computing Technology, Shenzhen University, China
[2] College of Computer and Cyber Security, Fujian Normal University, China
[3] Faculty of Information Science and Engineering, Ocean University of China, China
[4] Department of Computer and Information Science, Northumbria University, United Kingdom
[5] College of Computer Science and Software Engineering, Shenzhen University, China
[6] Guangdong Provincial Key Laboratory of Intelligent Information Processing, Shenzhen University, China
panyinghui@szu.edu.cn, zhaoshuaijie2023@email.szu.edu.cn, yushenbao@fjnu.edu.cn, zongyang.liu@stu.ouc.edu.cn, yifeng.zeng@northumbria.ac.uk, han.liu@szu.edu.cn, lmwfjnu@fjnu.edu.cn

Abstract

Consensus decision-making uses crowd responses (usually from non-experts) to questions to reach a consensus answer based on human-machine collaboration. The crucial point is dynamic, which should not only enable rapid self-iteration toward the correct answer through crowd workers' responses but also adaptively suggest the next most valuable question(s) to accelerate the integration of the answer. However, existing methods reach consensus using either offline data or fixed question search structures, thereby largely sidestepping this dynamic nature. In response, we propose a bilevel optimization-based human-machine collaboration (BiO-HMC), which explores an **inner & outer-level** optimization to enable effective answer integration and efficient question selection. The resulting optimization problem is intractable because there is no closed-form expression in the **inner-level** optimization. We employ a gradient-based method and guarantee the method's theoretical convergence. Experimental results on synthetic and real-world datasets demonstrate the effectiveness and efficiency of the BiO-HMC model, i.e., achieving the highest confidence in the correct answer with the lowest labor cost.

Code — https://github.com/zhaoshuaijie-cn/BiO-HMC

Introduction

With the rise of online platforms, consensus – a decision-making process in which participants work together to reach an answer – has embodied our desire to implement human-machine collaboration (Liu et al. 2015), where tasks are often difficult for computers to handle. In complex consensus tasks, because most crowd workers lack domain expertise, they are usually given a sequence of questions with minimal domain knowledge rather than being asked for the possible answers directly, and the correct answer is obtained by merging their opinions (Deng and Xiang 2021). Fig. 1 illustrates a crowdsourcing-based system that aids intuition, aiming to determine the galaxy of a given image. To solve this task, in Fig. 1(**b**), workers are presented sequentially with four questions, each including several options that describe the features the galaxy may have (see an example of the question-

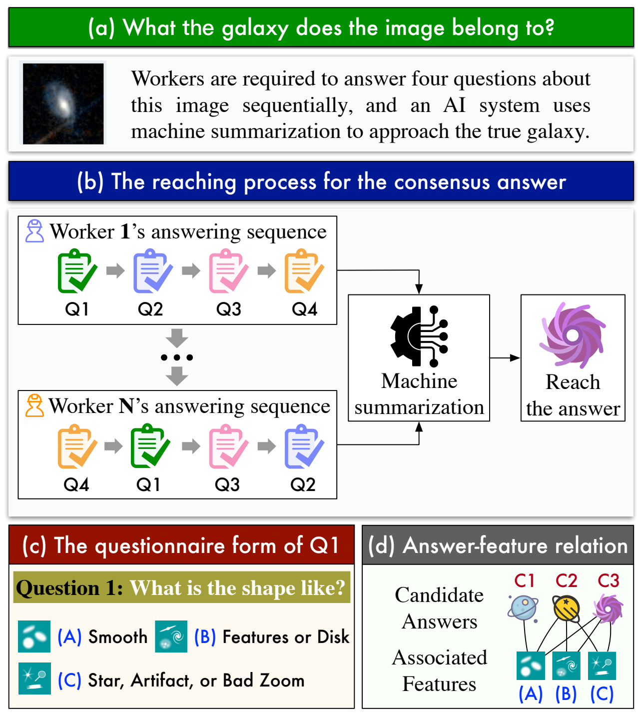


Figure 1: An illustration of multi-round crowdsourcing to determine the galaxy category of an image. (**a**) describes the task. In (**b**), we provide workers with questions (e.g., Q1 in (**c**)) about this image. The workers select the appropriate option to indicate whether a specific feature is present or absent. Note that the feature can support one or more candidate galaxy categories, see the relations in (**d**).

[*]The Corresponding Author

naire of Q1 in Fig. 1(**c**)). A worker's response to the question indicates whether a specific feature is present in this image. Note that the feature could support multiple categories, for example, the feature *smooth* of option '(A)' can support two potential galaxies ('No.1' and 'No.3' in Fig. 1(**d**)). After collecting their responses, the system uses machine summarization to derive a consensus answer (i.e., the true category of the galaxy), and the key lies in achieving the most accurate answer possible with the lowest labor costs.

To facilitate human-machine collaboration in such complex consensus tasks, a growing line of work is dedicated to the design of an accurate answering process (Siddharthan et al. 2016; Tu et al. 2020; Zhang, Jiang, and Li 2024), or endeavors to find planning strategies (Kamar and Horvitz 2013, 2015; Deng and Xiang 2021, 2025) that favor the expected answer. Despite varying degrees of success, these solutions primarily focus on answer integration, either using offline datasets or a fixed searching space of questions. These lazy strategies and fixed search structures not only limit the flexibility in question selection but also incur a computational burden on managing workers' responses, lacking further insight into the dynamic nature of the consensus. Recall the workers' answering process in Fig. 1(**b**), due to the uncertainty of their responses, this dynamic is by nature a multi-round decision-making process, and each round includes: (**i**) **a collection step** that requires updating the belief toward the correct answer, provided the workers' responses are gathered; (**ii**) **a planning step** that considers the relevance of questions to the true answer and can prioritize the valuable question(s) since low-relevance features could blur workers' judgment and increase labor burdens. Hence, the existing solutions are still far from perfect. There remains a need for an efficient method that can not only facilitate efficient answer integration but also suggest the next valuable guidance for workers without restriction. Unfortunately, the long collection & planning decision-making is NP-hard since it can be reduced to (generalized) troubleshooting problems with dependent actions and answer uncertainty (Vomlelová and Vomlel 2003).

To this end, our key observation is that bilevel programming (Colson, Marcotte, and Savard 2007) could enable the collection & planning steps, and we rethink this multi-round decision making as a bilevel optimization problem. However, unlike traditional bilevel programming adapted to supervised settings (Franceschi et al. 2018; Ghosh and Lan 2021), this formalization poses several challenges. The major issue is that the consensus task is unsupervised because the candidate answers are latent/unmeasured, and the true one is unknown. Although the probabilities (or say beliefs) of possible answers can be estimated from workers' responses to questions, a question may support multiple candidate answers, and responses are often noisy and limited because non-expert workers may only answer partial questions. Hence, the main hurdle is designing suitable objective functions for the closed-loop optimization.

To solve these challenges, we propose the BiO-HMC – a **Bi**level **O**ptimization-based **H**uman-**M**achine **C**ollaboration paradigm that facilitates an **inner & outer-level** optimization. In light of Bayesian networks (BN) that can seamlessly encode domain knowledge (Masegosa and Moral 2013), the **outer-level** optimization achieves the collection step by carefully designing a two-layer BN with the noisy-OR model (named NO_2). In NO_2, workers' responses are encoded as beliefs about candidate answers and the associated questions. These beliefs are optimized when new worker responses continuously arrive. In the **inner-level** optimization, we propose a question-prioritization model that selects the question with the highest value for the next optimization step. Hence, planning steps are achieved. As solving the bilevel optimization is difficult due to the lack of a closed-form solution in the **inner-level** optimization, which poses a dilemma for directly optimizing the outer objective function, we employ a gradient-based method and guarantee its theoretical convergence. As well as being of practical importance, the gradient method can not only alleviate computational burden but also enable real-time adjustment of the answers' beliefs, which helps to determine whether to terminate question collections. The main contributions are:

- We recast the consensus tasks as bilevel programming and propose the BiO-HMC framework, which enables closed-loop optimization to provide workers with the most valuable question(s), and the belief of the true answer can be discerned efficiently.
- We facilitate efficient bilevel optimization through (a) proposing NO_2 that taps into crowd responses to approach the correct answer; (b) designing a value preference-based function that prioritizes valuable question(s) to accelerate the consensus decision process. We also prove our bilevel method's theoretical convergence.
- We compare BiO-HMC with state-of-the-art planning algorithms on both synthetic and real-world datasets. Experimental results demonstrate the effectiveness and efficiency of our BiO-HMC, i.e., achieving the highest confidence in the correct answer with the lowest labor cost.

Related Works

A crowdsourcing consensus task involves two key components: (a) how to assign (sub)-tasks to crowdsourcing workers; and (b) how to aggregate workers' opinions to achieve high-quality answers (Pan 2016). To address these problems, massive efforts have been made over the past few decades. To name a few, for (a), Tran *et al.* (Tran-Thanh et al. 2014) employed a bounded multi-armed bandit algorithm, and Colovic *et al.* (Colovic, Bagherzadeh, and Liévin 2024) designed a crowdsourcing classification framework to achieve efficient task allocation. For (b), given that a multi-round questioning process could help improve the answer quality (Sheng, Provost, and Ipeirotis 2008), much research focuses on enhancing the questioning quality (Tu et al. 2020; Shu, Sheng, and Li 2018) to increase answer reliability, or designing suitable answer aggregation models (Zhang, Jiang, and Li 2024; Anagnostopoulos et al. 2015). In addition, other works appeared that explored the use of contextual information to refine the aggregation process (Chai, Sun, and Wang 2022; Zhu et al. 2022). However, these methods are primarily based on offline datasets, which cannot dynamically optimize decision-making, limiting their applicability to complex consensus tasks.

In another vein, researchers used partially observable Markov decision processes (POMDPs) to solve consensus tasks. For example, Kamar and Horvitz (Kamar and Horvitz 2013, 2015) developed a hierarchical classification model to integrate machine and human results across task layers. They employed Monte Carlo planning to calculate the expected value of information (VOI), thereby selecting actions with the highest VOI to accelerate consensus resolution. However, the reliance on predefined hierarchies limits its dynamic action adjustments. In addition, Deng and Xiang (Deng and Xiang 2021) introduced an option-candidate (OC) model with an improved depth-first algorithm to generate query strategies. The OC model enhances model flexibility but increases computational burdens. This idea is further pursued in (Deng and Xiang 2025) that applies a multi-agent system, in which each agent handles a search space to reduce running time, but is not intended to reduce labor interactions or improve answer accuracy.

The BiO-HMC Model

Given M questions $\mathcal{O} = \{O_1, \cdots, O_{\mathrm{M}}\}$ and K candidate answers $\mathcal{C} = \{C_1, \cdots, C_{\mathrm{K}}\}$, each question $O_m \in \mathcal{O}$ (or answer $C_k \in \mathcal{C}$) has a true-or-false binary value, where the true response ('$+$') of the question can support the relevant answers. The negative response ('$-$') considers the associated answers wrong. Note that one or more questions can support an answer, while multiple answers could share one question. Assuming that we collect N workers ($\mathbf{\Psi}=\{\psi_1, \cdots, \psi_{\mathrm{N}}\}$)' responses to (partial) questions sequentially, in which a response behavior is regarded as an interaction with a worker, our goal is to find the consensus answer as correctly as possible with the minimum number of interactions.

The Bilevel Solution

We start by sketching the BiO-HMC model to achieve the goal above. Given a worker ψ_n $(1 \leq n \leq \mathrm{N})$ with $t-1$ historical responses (or observations), denoted as $\Gamma^{(t-1)} = \{o^{(1)}, \cdots, o^{(t-1)}\}$ [1], the t-th response process can be formulated as solving the following bilevel optimization problem:

$$\theta^* \leftarrow \arg\max_{\theta} \ \ell(\theta) = \log P(\Gamma^{(t)}) \tag{1}$$

$$\begin{aligned} \text{s.t.} \quad & \pi_\theta(\mathcal{S}) \leftarrow \arg\max_{\mathcal{S}} \mathcal{L}(\mathcal{S}, \theta^{(t-1)}), \qquad (2) \\ & O^{(t)} \sim \pi_\theta(\Gamma^{(t-1)}; \mathcal{S}) \in \Omega^{(t)}, \\ & \Gamma^{(t)} = \Gamma^{(t-1)} \cup \{o^{(t)}\}. \end{aligned}$$

In problem (2), we select the most valuable question $O^{(t)}$ from the question bank $\Omega^{(t)}$ based on the proposed question prioritization algorithm $\pi_\theta(\cdot)$, and the worker's response $o^{(t)}$ contributes to $\Gamma^{(t)}$. $\pi_\theta(\cdot)$ is learned by maximizing the likelihood of value preference of questions, given the current $\theta^{(t-1)}$ estimated in the outer level. Here, $\mathcal{S} = \{s_1, \cdots, s_{\mathrm{M}}\}$ is a question value set, where $s_m \in \mathbb{R}$ denotes the value of O_m, and the higher the value of s_m, the better O_m contributes to determining the answer. In problem (1), we maximize the likelihood of the responses $\Gamma^{(t)}$ based on the proposed NO_2 model to estimate the beliefs of questions and answers, i.e., θ. After the closed-loop optimization of (1) and (2), we can move toward the correct answer. Next, we detail the **outer (inner)-level** objective functions.

[1]Since question sequence can be varied, we omit the subscript n of observation o when the context is clear.

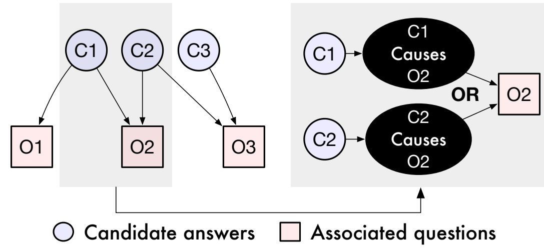


Figure 2: The NO_2 model based on two-layer BN. The nodes with circles (rectangles) are latent (observed) variables.

The outer-level objective function in problem (1). This optimization aims to update the beliefs for questions and answers (i.e., θ) using workers' historical responses through the proposed NO_2 model. Specifically, given the answer set $\mathcal{C}$ and the question set $\mathcal{O}$, we design the NO_2 model (see a schematic diagram in Fig. 2) to model $\mathcal{C}$ and $\mathcal{O}$ as latent and observed variables, respectively. Considering that an answer is characterized by its associated questions, meaning these questions contribute to forming the answer, we encode the causal relations between the question O_m and the candidate answer C_k via the noisy-OR gate. The marginal probability of C_k is $P(C_k)$, which quantifies the belief of the answer being true. The conditional probability of O_m is $P(O_m|\mathcal{C}_{\mathrm{pa}(m)})$, where $\mathcal{C}_{\mathrm{pa}(m)} \subseteq \mathcal{C}$ are O_m's latent parents. Given that the possible answers are mutually independent, namely $C_i \perp\!\!\!\perp C_j$ $(1 \leq i, j \leq \mathrm{K})$, and the presence of a question is individually caused by its associated answers, i.e., $O_i \perp\!\!\!\perp O_j | \mathcal{C}$ $(1 \leq i, j \leq \mathrm{M})$, a question O_m is false (say o_m^-) if and only if none of the answers in its latent parents $\mathcal{C}_{\mathrm{pa}(m)}$ cause O_m to be true, which gives

$$P\left(o_m^- | c_{\mathrm{pa}(m)}^+\right) = \prod_{c_k \in c_{\mathrm{pa}(m)}^+} P\left(o_m^- | \text{only } c_k^+\right),$$

where $P(o_m^- | \text{only } c_k^+)$ is the probability of O_m being false if only $C_k \in \mathcal{C}_{\mathrm{pa}(m)}$ is true (say c_k^+). Based on NO2, given N workers' responses $\mathbf{R} =< \Gamma_1, \cdots, \Gamma_{\mathrm{N}} >$, we maximize

$$\ell(\theta) = \log \prod_{n=1}^{\mathrm{N}} P\left(\Gamma_n\right) = \sum_{n=1}^{\mathrm{N}} \log P\left(\Gamma_n\right). \tag{3}$$

We now detail the probability $P\left(\Gamma_n\right)$ on Eq. (3). For ease of notation, we omit the subscript n and divide Γ into $\{\boldsymbol{o}^+, \boldsymbol{o}^-\}$, where $\boldsymbol{o}^+$ ($\boldsymbol{o}^-$) is the true (false) response set. Hence, $P(\Gamma) = P\left(\boldsymbol{o}^+, \boldsymbol{o}^-\right)$, which can be rewritten as (see the derivation details in the supplemental material):

$$P(\boldsymbol{o}^+, \boldsymbol{o}^- | \theta) = \sum_{\boldsymbol{o}' \in 2^{\boldsymbol{o}^+}} \left\{ (-1)^{|\boldsymbol{o}'|} \times \right. \tag{4}$$

$$\prod_{k=1}^{\mathrm{K}} \left[\left(\prod_{o \in \boldsymbol{o}' \cup \boldsymbol{o}^-} P(o^- | \text{only } c_k^+) \right) P(c_k^+) + P(c_k^-) \right] \Bigg\},$$

where $\theta = \{P(c_k), P(o_m|\text{only } c_k); 1 \le k \le \mathrm{K}, 1 \le m \le \mathrm{M}\}$. We adapt these beliefs, which can, in addition to monitoring the probability of the true answer, also serve as confidence levels of question values in the **inner-level** optimization, as described in the sequel.

The inner-level objective function in problem (2). This optimization focuses on prioritizing the most valuable question(s) for workers, instead of wasting labor on useless questions that can not help determine the consensus answer. However, evaluating a question's value is an open issue because the true answer is unknown in the questioning process. To tackle this problem, we define a (pairwise) value preference matrix (Volkovs and Zemel 2014) to measure the "value preference" of any question pair from an unsupervised learning perspective. Specifically, we first construct a worker ψ_n's value preference matrix $\mathbf{Y}^{\psi_n} \in [0|1]^{\mathrm{M}\times\mathrm{M}}$ based on the responses. $\mathbf{Y}_{ij}^{\psi_n}{=}1$ if ψ_n's response to question O_i to be true while O_j to be false, and we consider O_i is more valuable than O_j because O_i could support the true answer compared with O_j; otherwise, $\mathbf{Y}_{ij}^{\psi_n}{=}0$. Then, we introduce question values $\mathcal{S} = \{s_1, \cdots, s_{\mathrm{M}}\}$, where $s_i \in \mathcal{S}$ is the value of O_i that contributes to the true answer. Assuming that $\mathbf{Y}^{\psi_n}$ is an outcome of multiple draws from the value preference distribution, we have

$$P(\mathbf{Y}_{ij}^{\psi_n} | \mathcal{S}, \theta) \coloneqq \frac{e^{(\beta_i+\beta_j)(s_i - s_j)}}{\sum_{k \neq l} e^{(\beta_k + \beta_l)(s_k - s_l)}}, \tag{5}$$

where β_i denotes the confidence level of s_i, which measures the uncertainty associated with the question's value assigned to O_i. Note that the confidence level can be achieved from the beliefs θ in the **outer-level** optimization, as $\beta_i = \sum_{c_l} P(o_i = 1|\text{only } c_l) P(c_l)$.

Armed with Eq. (5), we arrive at maximizing the likelihood $\mathcal{L}(\mathcal{S}, \theta)$ (abbreviated as $\mathcal{L}$) as

$$\mathcal{L} = \log \prod_{n=1}^{\mathrm{N}} \prod_{i \neq j} \left(\frac{e^{(\beta_i+\beta_j)(s_i-s_j)}}{\sum\limits_{k \neq l} e^{(\beta_k+\beta_l)(s_k-s_l)}} \right)^{\mathbf{Y}_{ij}^{\psi_n}} + \frac{\rho}{2} \|\mathcal{S}\|_2^2.$$

Here, the ℓ_2-norm regularizer is included to aid in establishing the convergence results. By maximizing the likelihood, we can calculate the optimal values $\mathcal{S}^* = \{s_1^*, \cdots, s_{\mathrm{M}}^*\}$, which serve to prioritize the most valuable question(s).

Note that our model is built on binary question-answer relationships because workers' binary responses are typically binary in non-expert consensus tasks. However, for multiple-choice cases, we can either formulate them as binary ones or extend our BiO-HMC by replacing the **outer-level** noisy-OR with noisy-MAX. In addition, we could also use e.g., structural equation models to support continuous responses.

Model Optimization

In problems (1) and (2), we need to optimize two types of parameters: (a) the beliefs θ and (b) the question values $\mathcal{S}$. To learn the parameters, we employ a stochastic gradient-based algorithm (Bovey and Senalp 2012), and the derivation details are provided in our supplemental material.

Update rule of θ. The gradient step for the beliefs θ is $\theta^{(l+1)} = \theta^{(l)} + \eta_\theta^{(l)} \nabla \ell(\theta^{(l)})$, where $\eta_\theta^{(l)}$ is a suitable step size at the l-th iteration, and $\nabla \ell(\theta) \propto \frac{\nabla P(\boldsymbol{o}^+, \boldsymbol{o}^- | \theta)}{P(\boldsymbol{o}^+, \boldsymbol{o}^- | \theta)}$. The gradient of $P(\boldsymbol{o}^+, \boldsymbol{o}^-|\theta)$ w.r.t. $P(c_k)$ and $P(o_m|\text{only } c_k^+)$ are

$$\frac{\partial P(\boldsymbol{o}^+, \boldsymbol{o}^- | \theta)}{\partial P(c_k^+)} = \sum_{\boldsymbol{o}' \in 2^{\boldsymbol{o}^+}} \Bigg[(-1)^{|\boldsymbol{o}'|} \prod_{i \neq k} U_i(\boldsymbol{o}') \times \left(\prod_{o_m \in \boldsymbol{o}' \cup \boldsymbol{o}^-} P(o_m^- | \text{only } c_k^+) - 1 \right) \Bigg], \tag{6}$$

$$\frac{\partial P(\boldsymbol{o}^+, \boldsymbol{o}^- | \theta)}{\partial P(o_m^- | \text{only } c_k^+)} = \sum_{\substack{\boldsymbol{o}' \in \\ 2^{(\boldsymbol{o}^+ \setminus o_m)}}} \Bigg[(-1)^{|\boldsymbol{o}'|+1} \prod_{i \neq k} U_i(\boldsymbol{o}' \cup o_m) \times \left(\prod_{\substack{o_j \in \\ (\boldsymbol{o}' \cup \boldsymbol{o}^- \setminus o_m)}} P(o_j^- | \text{only } c_k^+) P(c_k^+) \right) \Bigg], \tag{7}$$

where $U_i(\boldsymbol{o}') = \left[\prod_{o_m \in \boldsymbol{o}' \cup \boldsymbol{o}^-} P(o_m^- | \text{only } c_i^+) \right] P(c_i^+) + P(c_i^-)$.

Update rule of $\mathcal{S}$. The gradient step for the element $s_m \in \mathcal{S}$ is $s_m^{(l+1)} \leftarrow s_m^{(l)} + \eta_s^{(l)} \frac{\partial \mathcal{L}}{\partial s_m}$, where

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial s_m} = & \sum_{n=1}^{\mathrm{N}} \left[\sum_{j \neq m} (\beta_m + \beta_j) \mathbf{Y}_{mj}^{\psi_n} - \sum_{i \neq m} (\beta_i + \beta_m) \mathbf{Y}_{im}^{\psi_n} \right] \\ & - \sum_{n=1}^{\mathrm{N}} \Bigg\{ \frac{\mathrm{C}^{\psi_n}}{\sum_{k \neq l} e^{\mathrm{D}_{kl}}} \Bigg[\sum_{l \neq m} (\beta_m + \beta_l) e^{\mathrm{D}_{ml}} \\ & - \sum_{k \neq m} (\beta_k + \beta_m) e^{\mathrm{D}_{km}} \Bigg] \Bigg\} + \rho s_m. \end{aligned} \tag{8}$$

Here, $\mathrm{C}^{\psi_n} = \sum_{i \neq j} \mathbf{Y}_{ij}^{\psi_n}$, and $\mathrm{D}_{kl} = (\beta_k + \beta_l)(s_k - s_l)$.

Accordingly, we summarize the decision process in Algorithm 1. We initialize θ and $\mathcal{S}$ with equal probabilities and define global and local value preference matrices ($\mathbf{Y}$ and $\mathbf{Y}'$). Ω is the question bank. For each iteration, we first empty $\mathbf{Y}'$ and add all questions into Ω (Line 3). After that, we select the most valuable question from Ω (Line 5). Then, we update θ (Lines 6-8). If the belief of any possible answer exceeds the confidence level λ, we return the answer as the final result (Line 9); otherwise, we continue the procedure and update $\mathcal{S}$ (Line 10). When all the questions are used up (i.e., $\Omega = \emptyset$), we merge $\mathbf{Y}'$ into $\mathbf{Y}$ to update the global value preferences and ask the next worker. The loop ends when the maximum number of workers is reached.

Theoretical Analysis

We now establish the convergence guarantee of Algorithm 1. Indeed, it is not easy to express the closed form of $\mathcal{S}$ due to the linear combination of the "log-sum-exp" expressions.

Algorithm 1: The Bilevel Optimization Algorithm

Input: The NO_2 model $\mathcal{B} = (\mathcal{O} \cup \mathcal{C}, \mathcal{E})$, the maximum worker numer N, the confidence level λ.
Output: The maximum value of $P(c_k^+)$ $(1 \leq k \leq \mathrm{K})$.

Initialize: (1) $P(c_k^+) \leftarrow 1/\mathrm{K}$; (2) $P(o_m^+|\text{only } c_k^+) \leftarrow 0.5$; (3) $\mathcal{S} \leftarrow 1/\mathrm{M}$; (4) the global and local value preference matrices $\mathbf{Y}$ and $\mathbf{Y}'$; (5) available question set Ω.
for $iter < \mathrm{N}$ **do**
  $\mathbf{Y}', n \leftarrow 0$; add $O_m (1 \leq m \leq \mathrm{M})$ into Ω;
  while $n < \mathrm{M}$ **do**
    Select O_m from Ω with the largest value in $\mathcal{S}$, and remove O_m from Ω and observe o_m;
    for (O_m, C_k) in $\mathcal{E}$ **do**
      Update $P(c_k)$ and $P(o_m|\text{only } c_k^+)$ by Eqs. (6) and (7), respectively;
    end for
    Return $P(c_k^+)$ if $P(c_k^+) > \lambda$;
    Update (i) $\mathbf{Y}'$ by o_m, (ii) β_i by $P(c_k)$ and $P(o_m|\text{only } c_k^+)$, and (iii) $\mathcal{S}$ by Eq. (8);
    $n \leftarrow n + 1$;
  end while
  Merge $\mathbf{Y}'$ into $\mathbf{Y}$;
end for
return The maximum value of $P(c_k^+)$ $(1 \leq k \leq \mathrm{K})$.

Hence, it is impossible to optimize the outer objective function directly. To solve this issue, we use the gradient-based method to achieve the optimum, and the following theorem establishes the convergence (Franceschi et al. 2018).

Theorem 1 *Let* $[\mathrm{T}] = \{1, 2, \cdots, \mathrm{T}\}$ *where* T *is a positive integer. If the outer-level* $\ell(\theta)$ *is uniformly Lipschitz continuous and the iterates* $\langle \mathcal{S}_\theta \rangle_{\mathrm{T} \in \mathbb{N}}$ *in the inner-level converge uniformly to* $\mathcal{S}_\theta$ *as* $\mathrm{T} \to +\infty$*, then* $\max \ell_{\mathrm{T}}(\theta) \to \max \ell(\theta)$*, meaning that for every sequence* $\langle \theta \rangle_{\mathrm{T} \in \mathbb{N}}$ *such that* $\theta \in \arg\max_\theta \ell_{\mathrm{T}}(\theta)$*, we have* $\langle \theta \rangle_{\mathrm{T}}$ *admits a convergent subsequence and for every subsequence* $\langle \theta_{sub} \rangle_{\mathrm{T}}$ *such that* $\theta_{sub} \to \bar{\theta}$*, where* $\bar{\theta} \in \max \ell(\theta)$.

To prove Theorem 1, the key lies in proving the **outer-level** $\ell(\theta)$ is *Lipschitz* continuous. The **inner-level** $\mathcal{L}$ can achieve a uniform solution given θ, i.e., $\mathcal{L}$ is concave. We prove that $\ell(\theta)$ is twice continuously differentiable with an upper bound, and discuss the concavity of $\mathcal{L}$. The proof details and the convergence analysis are available in our supplemental material.

Experimental Results

In this section, we evaluate the performance of BiO-HMC on four aspects: (**S1**) Can the BiO-HMC approach the correct answer with a good confidence level given unlimited interactions? (**S2**) How does BiO-HMC perform when the interactions are limited? (**S3**) How is the robustness of BiO-HMC when it comes to different noisy levels of workers' responses? (**S4**) How does the question selection algorithm affect the performance of correct answer identification? In addition to **S1**-**S4**, extended experiments can be found in the supplemental material.

Experiment Setup

Dataset description. We use five synthetic datasets (denoted as **Syncs1-5**) and a real-world public dataset from **Galaxy Zoo** (abbreviated as **GZ**) [2]. The question-answer relations of the synthetic datasets are summarized in the supplemental material. For each synthetic dataset, we randomly specify a correct answer and generate the binary values (with a probability of a predefined *noise rate* ζ) for the associated questions sequentially to simulate the workers' noisy responses. For GZ, it contains 45 candidate answers with 17,787 pieces of responses, and the answer-question relations are built based on the decision tree. Notably, GZ has served as the typical testbed for consensus studies, on which the evaluation results of models are highly persuasive.

Baseline approaches. We use the baseline approaches as follows: (**i**) **Limited lookahead** that exhaustively compares the beliefs over all questions with a limited lookahead depth. (**ii**) **MC-VOI** (Kamar and Horvitz 2015), a Monte-Carlo (MC) planning algorithm that uses VOI to evaluate the utility of any action outcome sequence. (**iii**) **MC-EVA** (Deng and Xiang 2021), a multi-step collection model based on the MC sampling with the expected value algorithm (EVA). (**iv**) **MC-EVA+AS** (Deng and Xiang 2021), an enhanced MC-EVA by the depth-first searching algorithm. (**v**) **PMC-Cnts** (Deng and Xiang 2025), which introduces a multi-agent system to improve MC-EVA. PMC-Cnts divides the question-searching space into multiple substructures.

Evaluation metrics. We use two metrics: **highest support rate** (**HSR**) and **Cost**. HSR considers the answer with the highest support from all workers to be true because a worker's responses to questions can support one or more answers (Kamar and Horvitz 2013). Note that we set the lower bound of HSR as 80%; otherwise, the reported answer is called *undecidable*. For Cost, we measure the cost of a method in terms of: (a) the number of interactions (i.e., the number of responses consumed), denoted as **# Int**, and (b) the **runtime** (seconds).

S1: Evaluation on Unlimited Responses

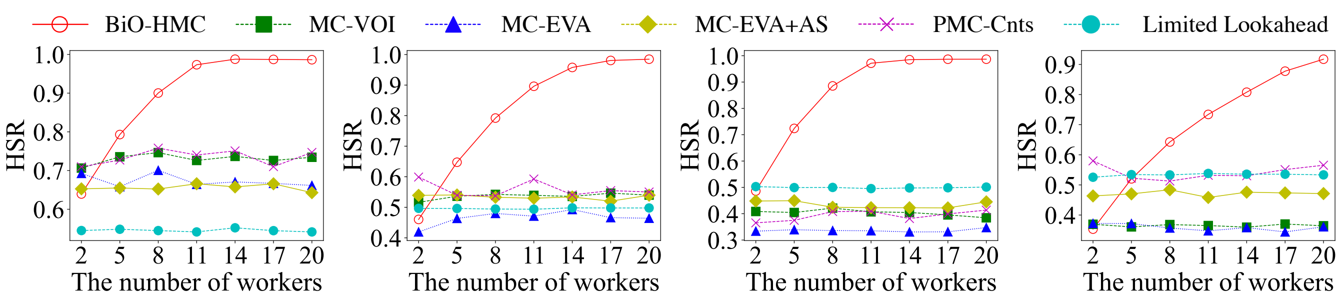
We first investigate the model performance under unlimited responses. For the synthetic datasets, the noise rate $\zeta = 20\%$ (we also discuss performance across different noise rates in the sequel). Table 1 provides the compared results, where we use boldface and cell shading to highlight the best performance. As shown in Table 1, BiO-HMC achieves the best performance in terms of all evaluation metrics across the synthetic datasets. While MC-VOI, MC-EVA, MC-EVA+AS, and PMC-Cnts perform well on Sync1 and Sync3, their HSR values decrease when dealing with the more complex answer-question relations on Sync2, Sync4, and Sync5. There are some numerical results below 80%, indicating that the corresponding methods (especially the limited lookahead) fail to reach consensus. In addition, PMC-Cnts reduces MC-EVA's runtime by introducing multiple agents. However, PMC-Cnts introduces shared observations, increasing the number of interactions. Instead, BiO-HMC

[2] https://data.galaxyzoo.org/

Dataset	Metric	MC-VOI		MC-EVA		MC-EVA+AS		PMC-Cnts		Limited lookahead	BiO-HMC
Sync1	HSR ↑		0.987		0.986		0.985		**0.988**	0.724*	**0.988**
	# Int ↓		466		450		1,430		1,990	600	62
	runtime ↓		0.038		0.069		0.183		0.035	6.372	0.029
Sync2	HSR ↑		0.902		0.849		0.898		0.936	0.592*	**0.987**
	# Int ↓		8,010		8,357		6,053		14,431	600	84
	runtime ↓		2.214		1.798		0.965		0.752	41.957	0.205
Sync3	HSR ↑		0.978		0.951		0.949		0.985	0.655*	**0.986**
	# Int ↓		2,873		4,113		4,777		11,428	600	141
	runtime ↓		1.149		0.539		0.859		0.226	32.763	0.187
Sync4	HSR ↑		0.761*		0.692*		0.728*		0.810	0.665*	**0.987**
	# Int ↓		9,668		9,671		9,340		13,937	600	123
	runtime ↓		5.650		1.765		1.854		0.996	179.916	0.541
Sync5	HSR ↑		0.887		0.799*		0.844		0.925	0.655*	**0.986**
	# Int ↓		7,678		9,340		8,020		13,768	600	131
	runtime ↓		2.416		1.538		2.221		1.168	104.443	0.321
Dataset	Metric	100^{+}	200	100	200	100	200	100	200	\	\
GZ	HSR ↑	0.791*	0.955	0.680*	0.879	0.709*	0.899	0.793*	0.928	0.603*	**0.983**
	# Int ↓	9,750	9,110	9,670	10,187	9,340	12,740	12,501	25,118	600	287
	runtime ↓	8.381	4.356	3.517	4.412	3.262	4.892	2.667	3.993	359.696	2.747

* The HSR is lower than 80%, which means the correct answer is *undecidable*.
$^{+}$ Here, 100 denotes MC-VOI with *horizon*=100. 200 can be explained analogously.

Table 1: The compared results of BiO-HMC with baselines with unlimited responses



(a) Sync1 (b) Sync3 (c) Sync5 (d) Galaxy Zoo (GZ)

Figure 3: The comparison of BiO-HMC with baselines in terms of HSR using limited responses.

requires the fewest interactions and has the shortest runtime, suggesting its high efficiency.

For GZ, while MC-VOI, MC-EVA (+AS), and PMC-Cnts have high support rates, these methods incur higher costs due to their long-horizon values, which reduces the efficiency of reaching consensus. For example, MC-EVA+AS with *horizon* = 200 outperforms the one with *horizon* = 100 (0.899 *vs.* 0.709), but the former requires more interactions and a longer runtime. We also observe that the horizon has a limited effect, so continuously increasing it does not improve the support rate. In contrast, BiO-HMC achieves a good support rate while providing large cost savings.

S2: Evaluation on Limited Responses

Next, we evaluate the efficiency of BiO-HMC in finding the consensus answer given limited responses. Considering that a large number of candidate answers and questions require more workers while smaller ones do not, we set the maximum number of workers to 20, and the noise rate ζ remains 20%. Accordingly, we set the horizon for MC-VOI, MC-EVA, MC-EVA+AS, and PMC-Cnts to 30. Fig. 3 summarizes the compared results on Syncs 1, 3, 5, and GZ, and the full visualization refers to the supplemental material. As shown in Fig. 3, the HSR value of BiO-HMC improves when we increase the number of workers, which aligns with expectations that our method can obtain the correct answer during the multi-round decision process. Notably, on the synthetic datasets, BiO-HMC achieves 80% accuracy after interacting with approximately 8 workers, demonstrating its efficiency. In contrast, the baseline methods show stable (or slightly increasing) HSR values, with most remaining below 80%, indicating that the baselines require more interactions

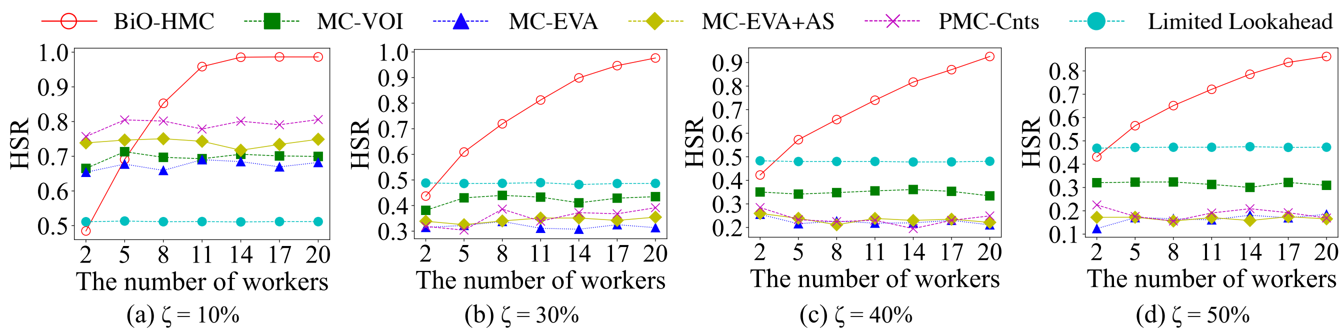


Figure 4: The compared results (HSR) of BiO-HMC and baselines using limited responses with different noise rates on Sync3.

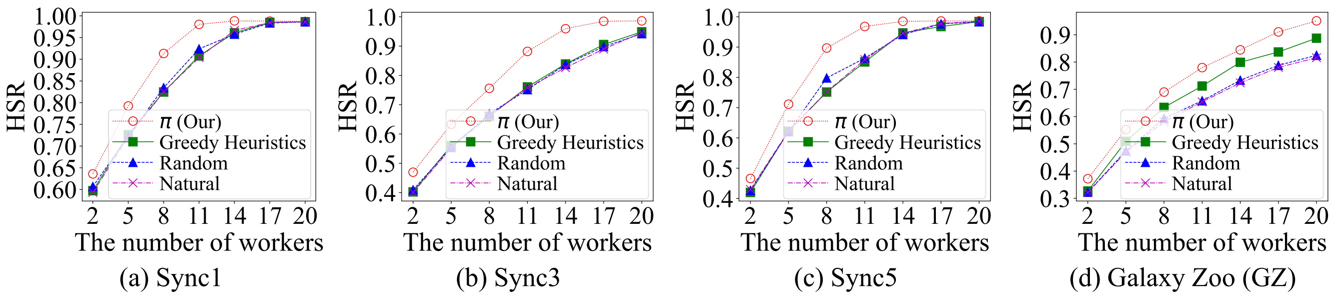


Figure 5: Effects of different question selection strategies on the performance of HSR ($\zeta = 20\%$).

to increase the confidence in the correct answer.

S3: Robustness Analysis

Recall that we introduced the noise rate ζ to account for noisy observations, and a large ζ value may blur the consensus answer. In this section, we discuss the robustness of BiO-HMC compared with baselines across different noise rates in the limited-responses case. Since we compared the performance based on $\zeta = 20\%$ in **S2**, we show the results using $\{10\%, 30\%, 40\%, 50\%\}$ on Sync3 in Fig. 4, and the results on the remaining datasets refer to the supplemental material. We observe that when $\zeta = 10\%$, BiO-HMC begins to fall behind several baselines. However, as we increase the number of workers, our method can rapidly improve HSR and outperform all the compared methods after a few interactions. Notably, BiO-HMC achieves similar HSR values (all above 98%) across different noise rates when finishing the decision process, demonstrating its robustness. In contrast, most baselines exhibit fluctuating performance with only slight changes in HSR (similar to the results in **S2**), indicating their higher sensitivity to noise.

S4: Question Selection Comparison

In BiO-HMC, we design $\pi_\theta(\cdot)$ to prioritize questions and accelerate the search for the correct answer. To demonstrate its effectiveness, we compare it with other selection strategies, including *natural*, *random*, and *greedy heuristics*. The natural strategy ranks questions based on the ordinal number, while the random strategy shuffles the questions randomly for each new worker. The greedy heuristics selects the most valuable questions based on their conditional beliefs. We report the compared results (with ζ=20%) on Syncs 1, 3, 5, and GZ in Fig. 5 (full results refer to our supplemental material). As observed, $\pi_\theta(\cdot)$ outperforms the comparison methods across all datasets. For the datasets with simple answer-question relations (e.g., Sync1), our algorithm achieves a support rate of over 98% with several interactions. For the more complex relations (e.g., GZ), our selection algorithm also shows higher data efficiency. Therefore, Fig. 5 validates the positive performance of the proposed question selection algorithm, confirming its efficiency.

Conclusion

In this paper, we investigate the consensus task of finding the consensus answer using noisy responses to the associated questions from non-expert crowdsourcing workers. To this end, we propose the BiO-HMC model, which employs a bilevel optimization framework to enable effective answer integration and efficient question prioritization, where the former is achieved through the proposed NO_2 model and the latter is approached by the careful design of value preferences for questions. Experiments on both synthetic and real-world datasets demonstrate the effectiveness and efficiency of our model, i.e., finding the consensus answer as accurately as possible with minimal human-machine interaction, thereby saving labor costs. The future work is two-fold: (1) considering the skill levels of workers if available, and (2) extending our model to discern the *unknown unknowns* when the training data does not include the correct answer during human-machine interaction.

Acknowledgments

Yinghui Pan was supported by the National Natural Science Foundation of China (Grant No. 62276168) and the Scientific Foundation for Youth Scholars of Shenzhen University, China (Grant No. 868-000001032177). **Shenbao Yu** was supported by the National Natural Science Foundation of China (Grant No. 62506078), the Science Foundation for Youth of the Education Department of Fujian Province, China (Grant No. JZ240013), and the Natural Science Foundation of Fujian Province, China (Grant No. 2025J08162). **Han Liu** was supported by the National Natural Science Foundation of China (Grant 62106147), Shandong Provincial Natural Science Foundation, China (Grant No. ZR2025QC1591), Guangdong Provincial Key Laboratory, China (Grant No. 2023B1212060076), Shenzhen Science and Technology Program, China (Grant No. ZDSYS20220527171400002), and Scientific Foundation for Youth Scholars of Shenzhen University, China (Grant 827-000798).

References

Anagnostopoulos, A.; Becchetti, L.; Fazzone, A.; Mele, I.; and Riondato, M. 2015. The importance of being expert: efficient max-finding in crowdsourcing. In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, 983–998.

Bovey, R. L.; and Senalp, E. 2012. Suggesting and justifying model updates for improved troubleshooting. *Quality and Reliability Engineering International*, 28(6): 616–623.

Chai, L.; Sun, H.; and Wang, Z. 2022. An error consistency based approach to answer aggregation in open-ended crowdsourcing. *Information Sciences*, 608: 1029–1044.

Colovic, A.; Bagherzadeh, M.; and Liévin, J.-L. 2024. Choosing the right crowdsourcing strategy: implications for governments' crowdsourcing initiatives. *Public Administration Review*.

Colson, B.; Marcotte, P.; and Savard, G. 2007. An overview of bilevel optimization. *Annals of Operations Research*, 153: 235–256.

Deng, Z.; and Xiang, Y. 2021. Multistep planning for crowdsourcing complex consensus tasks. *Knowledge-Based Systems*, 231: 107447.

Deng, Z.; and Xiang, Y. 2025. A partitioning monte carlo approach for consensus tasks in crowdsourcing. *Expert Systems with Applications*, 262: 125559.

Franceschi, L.; Frasconi, P.; Salzo, S.; Grazzi, R.; and Pontil, M. 2018. Bilevel programming for hyperparameter optimization and meta-learning. In *Proceedings of the 35th International Conference on Machine Learning*, 1568–1577.

Ghosh, A.; and Lan, A. 2021. bobcat: bilevel optimization-based computerized adaptive Testing. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence*, 2410–2417.

Kamar, E.; and Horvitz, E. 2013. Light at the end of the tunnel: a Monte Carlo approach to computing value of information. In *International Conference on Autonomous Agents and Multi-Agent Systems*, 571–578.

Kamar, E.; and Horvitz, E. 2015. Planning for crowdsourcing hierarchical tasks. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, 1191–1199.

Liu, S.; Miao, C.; Liu, Y.; Yu, H.; Zhang, J.; and Leung, C. 2015. An incentive mechanism to elicit truthful opinions for crowdsourced multiple choice consensus tasks. In *International Conference on Web Intelligence and Intelligent Agent Technology*, 96–103.

Masegosa, A. R.; and Moral, S. 2013. An interactive approach for Bayesian network learning using domain/expert knowledge. *International Journal of Approximate Reasoning*, 54(8): 1168–1181.

Pan, S. 2016. *Dynamic crowdsourcing consensus tasks with workers that can learn*. Master's thesis, University of Waterloo.

Sheng, V. S.; Provost, F.; and Ipeirotis, P. G. 2008. Get another label? improving data quality and data mining using multiple, noisy labelers. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 614–622.

Shu, Z.; Sheng, V. S.; and Li, J. 2018. Learning from crowds with active learning and self-healing. *Neural Computing and Applications*, 30(9): 2883–2894.

Siddharthan, A.; Lambin, C.; Robinson, A.-M.; Sharma, N.; Comont, R.; O'mahony, E.; Mellish, C.; and Wal, R. V. D. 2016. Crowdsourcing without a crowd: reliable online species identification using Bayesian models to minimize crowd size. *ACM Transactions on Intelligent Systems and Technology*, 7(4): 1–20.

Tran-Thanh, L.; Stein, S.; Rogers, A.; and Jennings, N. R. 2014. Efficient crowdsourcing of unknown experts using bounded multi-armed bandits. *Artificial Intelligence*, 214: 89–111.

Tu, J.; Yu, G.; Domeniconi, C.; Wang, J.; Xiao, G.; and Guo, M. 2020. Multi-label crowd consensus via joint matrix factorization. *Knowledge and Information Systems*, 62: 1341–1369.

Volkovs, M. N.; and Zemel, R. S. 2014. New learning methods for supervised and unsupervised preference aggregation. *The Journal of Machine Learning Research*, 15(1): 1135–1176.

Vomlelová, M.; and Vomlel, J. 2003. Troubleshooting: np-hardness and solution methods. *Soft Computing*, 7(5): 357–368.

Zhang, Y.; Jiang, L.; and Li, C. 2024. Instance redistribution-based label integration for crowdsourcing. *Information Sciences*, 674: 120702.

Zhu, P.; Wang, Z.; Hauff, C.; Yang, J.; and Anand, A. 2022. Answer quality aware aggregation for extractive qa crowdsourcing. In *Findings of the Association for Computational Linguistics: EMNLP 2022*, 6147–6159.