FAIR: Fair Collaborative Active Learning with Individual Rationality for Scientific Discovery

Xinyi Xu[1 3] **Zhaoxuan Wu**[1 2] **Arun Verma**[1] **Chuan Sheng Foo**[3 4] **Bryan Kian Hsiang Low**[1]
[1]Department of Computer Science, National University of Singapore
[2]Institute of Data Science & ISEP NUSGS, National University of Singapore
[3]Institute for Infocomm Research (I^{2}R), Agency for Science, Technology and Research (A*STAR), Singapore
[4]Centre for Frontier AI Research (CFAR), Agency for Science, Technology and Research (A*STAR), Singapore

Abstract

Scientific discovery aims to find new patterns and test specific hypotheses by analysing large-scale experimental data. However, various practical limitations (e.g., high experimental costs or the inability to perform some experiments) make it challenging for researchers to collect sufficient experimental data for successful scientific discovery. To this end, we propose the *collaborative active learning* (CAL) framework that enables researchers to share their experimental data for mutual benefit. Specifically, our proposed coordinated acquisition function sets out to achieve *individual rationality* and *fairness* so that everyone can equitably benefit from collaboration. We empirically demonstrate that our method outperforms existing batch active learning ones (adapted to the CAL setting) in terms of both learning performance and fairness on various real-world scientific discovery datasets (biochemistry, material science, and physics).

1 INTRODUCTION

Scientific discoveries in biochemistry (Borkowski et al., 2020; Kabir and Wong, 2022), material science (Bassman et al., 2018), physics (de Silva et al., 2020), and other areas rely on performing experiments to collect large-scale observational data to find patterns and test specific hypotheses. However, many practical limitations (e.g., high cost of performing experiments[1] or the inability to do some experiments) make it fundamentally challenging for researchers to gather sufficient experimental data by themselves for successful scientific discovery. To overcome these limitations, we propose a collaboration framework that allows researchers to collaborate and share their experimental data for mutual benefit.

These collaborations among researchers lead to two main benefits. (1) *Cost saving*: Performing some experiments can be costly, so researchers can share their experimental data to reduce costs. For example, synthesising new drugs is expensive due to the cost of chemicals and the long experimentation time, so biochemists can reduce the costs of experiments by sharing their data. (2) *Principled quid-pro-quo*: If the researchers want to collect specific experimental data that are unobtainable by themselves but accessible to others, then through collaboration, the researchers can exchange their data to overcome such limitations. For example, astrophysicists rely on observed trajectories of planets to discover the differential equation governing a planetary system. However, constrained by their geographical locations, they cannot observe the whole trajectory by themselves (Voiland, 2017; Dunn, 2022). With collaboration (i.e., data sharing) among astrophysicists located at different geographical locations, they all can accurately learn about the underlying differential equation.

However, in practice, the norm is sharing the analysis obtained from the experimental data but *not* the data *per se* (Rousseau and Ustyuzhanin, 2022). It is because data sharing can lead to issues such as misperforming secondary analyses due to a lack of familiarity of the data collection process (Longo and Drazen, 2016), the original researchers not receiving appropriate credit for collecting the data (Park and Greene, 2018), or other researchers publishing the findings before the researchers who collected data (Taichman et al., 2016; Emmert-Streib et al., 2016). These issues can contribute toward a zero-sum paradigm as sharing the collected data can prove detrimental to those who collected the data (Lo and DeMets, 2016) and discourage

[1]For example, the high cost can arise from running highly specialised, expensive equipment like the Large Hadron Collider at CERN or quantum computers with high maintenance costs.

Proceedings of the 26th International Conference on Artificial Intelligence and Statistics (AISTATS) 2023, Valencia, Spain. PMLR: Volume 206.

collaboration, thus hindering research progress.

One possibility to address these issues is to enable researchers to collaborate during data collection, promote familiarity with the data, form shared authorship, and agree on the priority of research findings (Lo and DeMets, 2016). Then, sharing the data collected through such collaborations will provide the benefits of cost-saving and principled quid-pro-quo. Specifically, we consider *active learning* (AL) as the data collection framework because it is a principled experimental design method that utilises previous data points to efficiently select subsequent data points and is often used in biochemistry (Kell, 2012; Schneider, 2018), material science (Bassman et al., 2018; Gubaev et al., 2019; Lookman et al., 2019), and physics (Atkinson et al., 2019).

This paper proposes the *collaborative active learning* (CAL) framework that enables agents (researchers) to collaborate during data collection and requires each agent to specify its *support* (i.e., what data it can collect) and *target* (i.e., what data it wants). To encourage the collaboration, CAL should satisfy *individual rationality* (IR) (Sim et al., 2020; Tay et al., 2022) and *fairness* (Li et al., 2020, 2021; Sim et al., 2021). Here, IR ensures that collaboration benefits everyone, and fairness ensures that benefits are equitable. While existing batch AL methods may seem promising, they are not guaranteed to satisfy IR and fairness as they do *not* explicitly account for different agents who may have different supports and targets (Chen and Krause, 2013).

The first question we face is *what would be a suitable form of collaboration*. To answer this, we propose a coordinated acquisition function that enables the agents to leverage each other's support to learn about their targets. With this, we can characterise the benefit to an agent in collaboration via (a *lower* bound of) the *information gain* (IG) on its target. Higher IG implies that the agent can learn about its target effectively. Importantly, IR is satisfied (i.e., the agent is better off) if the IG for an agent with collaboration is at least what the agent can achieve individually.

It naturally raises the question of *how to characterise what an agent can achieve without collaboration*. What an agent can learn individually about its target would be limited if the agent's support cannot "cover" the target due to its observation constraints. Specifically, we derive an IG *upper* bound of the agent's target based on a distance-based formalisation of its observational constraints. The IG upper bound will be lower if the support is further from its target. Then, we derive a sufficient condition for IR, that is, the IG lower bound with collaboration is higher than the IG upper bound without collaboration.

The other key aspect of CAL is fairness which ensures that the agents benefit equitably to avoid unfair and exploitative collaborations (Li et al., 2020, 2021; Sim et al., 2021). It naturally leads to the question of *how to achieve fairness in CAL*. Intuitively, during data collection, if the IGs of all agents increase at an equitable rate, then their final IGs are also equitable, which entails fairness. To this end, we adopt a modified Nash social welfare (Kaneko and Nakamura, 1979) to show that our proposed coordinated acquisition function, which carefully selects what data to collect, effectively achieves fairness.

Our specific contributions of this work are as follows:

- We propose FAIR: Fair collaborative Active learning with Individual Rationality via a coordinated acquisition function to enable agents to collaborate during data collection (Sec. 3);
- We derive the IG upper bound of an agent without collaboration using the observational constraint between its support and target, then derive the IG lower bound of the agent with collaboration, and combine both results to characterise the sufficient condition for IR (Sec. 3.1);
- We exploit a modified Nash social welfare to show that the proposed coordinated acquisition function satisfies a generalised Pigou Dalton Principle, which effectively achieves fairness (Sec. 3.2); and
- We perform extensive experiments on synthetic data and real-world scientific discovery datasets in biochemistry, material science and physics. We demonstrate that our method outperforms existing batch AL ones in achieving better performance (i.e., IR) and fairness (Sec. 4).

2 PRELIMINARIES

Gaussian process (GP) in information-theoretic AL. A GP is defined by a mean function $\mu : \mathcal{X} \mapsto \mathbb{R}$ and a kernel function $\mathcal{K} : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}$ (e.g., squared exponential) where $\mathcal{X}$ is the input space for the data. Denote the ground truth function as $f : \mathcal{X} \mapsto \mathbb{R}$ (i.e., a regression setting). Moreover, the observed regression response is modelled with an additive Gaussian noise: $y := f(x) + \mathcal{N}(0, \lambda)$ (Rasmussen and Williams, 2006). Though the noisy observations y are *not* explicitly used in AL, this modelling choice affects the calculation of *information gain* (IG). Denote the acquired data as $X_m := \{x_1, \ldots, x_m\}$ such that each data/input location is from a *support* $\mathcal{S} \subseteq \mathcal{X}$. Then, denote the kernel-based gram matrix $K_{X_m} \in \mathbb{R}^{m \times m}$ on X_m as $K_{X_m} := [\mathcal{K}(x, x')]_{x,x' \in X_m}$ and a noise variance-regularised $\tilde{K}_{X_m} := K_{X_m} + \lambda \mathbf{I}$.

Denote the target $\mathcal{T} := \{x^*_1, \ldots, x^*_{m'}\} \subseteq \mathcal{X}$ as the data/input locations to learn about. We assume a conditional independence between $\mathcal{S}, \mathcal{T}$: $\forall X_A, X_B \subset \mathcal{S}$, s.t. $X_A \cap X_B = \emptyset, X_A \perp X_B | \mathcal{T}$. Using GP, we can place a covariance (matrix) over $\mathcal{T}$: denote $K_\mathcal{T} := \Sigma_\mathcal{T}$ as the prior covariance over $\mathcal{T}$ (i.e., before acquiring X_m); denote $\Sigma_{\mathcal{T}|X_m} := K_\mathcal{T} - K_{\mathcal{T}X_m} \tilde{K}^{-1}_{X_m} K_{X_m \mathcal{T}}$ as the posterior covariance over $\mathcal{T}$ (i.e., after acquiring X_m) where $K_\mathcal{T} \in \mathbb{R}^{m' \times m'}$ is the gram matrix of $\mathcal{T}$ (i.e., $K_\mathcal{T}[p, q] = \mathcal{K}(x^*_p, x^*_q)$) and $K_{\mathcal{T}X_m} = K^\top_{X_m \mathcal{T}} \in \mathbb{R}^{m' \times m}$ is

the covariance between $\mathcal{T}$ and X_m such that $K_{\mathcal{T}X_m}[p,q] := \mathcal{K}(x_p^*, x_q)$ between the p-th target data and q-th acquired data. The covariances $K_{\mathcal{T}}$ and $\Sigma_{\mathcal{T}|X_m}$ represent the uncertainty, respectively, before and after acquiring the data X_m through differential entropy $\mathbb{H}[\mathcal{T}|X_m]$, which quantifies the uncertainty about the target $\mathcal{T}$ conditioned on the acquired data X_m.[2] Based on this, information-theoretic AL methods (e.g., (Krause et al., 2008)) adopt the view of minimising uncertainty via the mutual information $\mathbb{I}[\mathcal{T}; X_m] := \mathbb{H}[\mathcal{T}] - \mathbb{H}[\mathcal{T}|X_m]$. Consequently, IG is defined as $\text{IG}(\{x\}; X_m, \mathcal{T}) := \mathbb{I}[\mathcal{T}; X_m \cup \{x\}] - \mathbb{I}[\mathcal{T}; X_m]$, with the following closed-form expression in GP:

$$\text{IG}(\mathcal{T}; X_m) = 0.5(\log\det\Sigma_{\mathcal{T}} - \log\det\Sigma_{\mathcal{T}|X_m})\,. \quad (1)$$

However, the exact maximisation of IG poses an NP-hard subset selection problem with a budget to collect m data points. To address this, the submodularity of IG is exploited to show that greedily collected X_m over m iterations satisfy near-optimality: $\text{IG}(\mathcal{T}; X_m) \geq (1-1/e)\,\text{IG}(\mathcal{T}; X^*)$ (Nemhauser et al., 1978). The conditional independence assumption mentioned above between (disjoint subsets of $\mathcal{S}$) and $\mathcal{T}$ is used to ensure the submodularity of Eq. (1) (Krause and Guestrin, 2005, Corollary 4). A further discussion is provided in Appendix A.

In addition to admitting such a near-optimal approximation, IG is often used as a surrogate measure for learning effectiveness due to a lack of knowledge of the function space (the reason for AL) (Kirsch et al., 2019; Chen and Krause, 2013). In contrast, in machine learning, the accuracy of a test set is often used to evaluate the learning performance. However, in scientific discoveries, due to no knowledge of search space, such a test set is unavailable, so IG lends itself to be a useful measure (evaluated in our experiments in Sec. 4).

3 FAIR: FAIR COLLABORATIVE ACTIVE LEARNING WITH INDIVIDUAL RATIONALITY

We propose FAIR, specified by a coordinated acquisition function, derive the conditions for *individual rationality* (Sec. 3.1) and provide the fairness guarantee (Sec. 3.2). Due to the page limitation, we defer the derivations and proofs, along with an overview of the key assumptions used in presented results to Appendix A.

Coordinated acquisition function. Denote n agents as $N := \{1, \ldots, n\}$ and $\mathcal{S}_i, \mathcal{T}_i \subseteq \mathcal{X}$ are the respective *support* and *target* of an agent $i \in N$. Let $\mathcal{S}_N := \bigtimes_{i\in N} \mathcal{S}_i$ be the Cartesian product of the supports of all the agents. Denote $\vec{x}_{N,k}$ (resp. $X_{N,k}$) as the data acquired by all agents in iteration k *only* (resp. up to and including iteration k). Define the shorthand $\text{IG}_{i,k} := \text{IG}(\mathcal{T}_i; X_{N,k})$ and $\Delta_{i,k} := \text{IG}_{i,k} - \text{IG}_{i,k-1}$. For simplicity, assume all agents have an equal budget of m. In iteration k, agent i acquires one data point from its support $\mathcal{S}_i$, so there are m iterations of the following coordinated acquisition function:

$$X_{N,k} \leftarrow X_{N,k-1} \cup \operatorname*{argmax}_{\vec{x}_{N,k}\in\mathcal{S}_N} \textstyle\sum_i \Delta_{i,k}/\beta_i, \quad (2)$$

where $\beta_i > 1$ are the *sharing coefficients* and w.l.o.g. $\sum_i 1/\beta_i = 1$. A larger β_i means the agent i is more willing to share, as $\text{IG}_{i,k}$ is maximised with a lower priority (i.e., $1/\beta_i$). More details on the effects of and how to set β_i are in Sec. 3.2 and Sec. 4. The specific optimisation objective is to leverage a near-optimality result (Krause et al., 2008) to provide a performance guarantee (Lemma 6 in Appendix A) and to derive IR (Sec. 3.1).

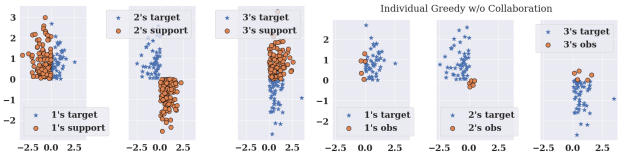

Figure 1: Left: Mismatch between $\mathcal{T}_i, \mathcal{S}_i$ and match between $\mathcal{T}_i, \mathcal{S}_j$ for $j \neq i$ imply each agent can benefit significantly from collaboration. Right: Individually, each agent's support (orange circles) cannot cover their targets (blue stars), so the learning effectiveness (i.e., IGs) is low.

To illustrate how Eq. (2) provides the quid-pro-quo benefit, if $\mathcal{S}_i$ is far from $\mathcal{T}_i$, but there is some other $\mathcal{S}_j$ which is close to $\mathcal{T}_i$, then letting agent j acquire a data point from $\mathcal{S}_j$ near $\mathcal{T}_i$ is beneficial to agent i (about $\mathcal{T}_i$) that will reciprocate to agent j or some other agent. For instance, in Fig. 1, with 3 agents, the support $\mathcal{S}_1$ of agent 1 cannot effectively cover its own target $\mathcal{T}_1$, but can help cover agent 2's target $\mathcal{T}_2$. If these agents do not collaborate, then Fig. 1 (right) shows that none of them can effectively cover their own target to learn effectively. Intuitively, since their supports can cover someone else's target, if these agents collaborate (i.e., adopt Eq. (2)), then their learning effectiveness can all improve (empirically verified later in Sec. 4, see Fig. 2).

One alternative to Eq. (2) is to explicitly instruct the agents to acquire data points to help each other. However, it is not guaranteed that the target of an agent is close to the support of another, which means extra care must be taken to ensure that no agent acquires data points outside its support. Moreover, it can be difficult to determine which agent should help which other agent. In contrast, formulating the acquisition function as an optimisation over the Cartesian product $\mathcal{S}_N$ avoids these difficulties. The cost-saving benefit can be seen as follows, suppose all the targets and supports are close (all the agents are interested in similar targets and can observe them closely), but it requires a high budget (costly for each agent individually). Then, Eq. (2) enables

[2] A set of data (e.g., $X_m \subseteq \mathcal{X}$ or $\mathcal{T} \subseteq \mathcal{X}$) are treated as a collection of random variables at the corresponding locations of the data, for $\mathbb{H}$ and subsequently $\mathbb{I}$ to be well defined.

the agents to pool their budgets to learn about their targets in a coordinated way. For instance, based on Eq. (2), two agents are unlikely to pick the same data point, as doing so is unlikely optimal (Sim et al., 2020).

3.1 Individual Rationality (IR)

Collaboration needs to result in a higher (or equal) IG for an agent than without collaboration to satisfy the IR. We first characterise the limitation to learning without collaboration (via an IG upper bound) due to observational constraints (i.e., $\mathcal{S}_i$ is far from $\mathcal{T}_i$). Next, we derive the conditions for collaboration to result in IG for i being higher than the upper bound without collaboration, hence achieving IR.

3.1.1 Limitation to Learning Without Collaboration

As the analysis is w.r.t. an agent i, the subscript i is omitted for simplicity: $\mathcal{T} \coloneqq \mathcal{T}_i$, $\mathcal{S} = \mathcal{S}_i$. Let $X_m \subseteq \mathcal{S}$ denote the acquired data by i *without* collaboration. By utilising a distance-based observational constraint (between $\mathcal{S}$ and $\mathcal{T}$), Proposition 1 gives the IG upper bound w.r.t. a singleton $\mathcal{T}$.

Proposition 1. For singleton $\mathcal{T} = \{x^*\}$, v_* is the prior variance at x^*, and budget m. Suppose $\inf_{x\in\mathcal{S},x^*\in\mathcal{T}} \|x - x^*\|_2 \geq \delta$. Let

$$M_* \coloneqq \left[\mathrm{tr}(\widetilde{K}_{X_m}) - (m-1)\left(\frac{m\det(\widetilde{K}_{X_m})}{\mathrm{tr}(\widetilde{K}_{X_m})}\right)^{\frac{1}{m-1}}\right] m\,\varepsilon_{\mathcal{K}}^2$$

where $\varepsilon_{\mathcal{K}} \coloneqq \sup_{x\in\mathcal{S}} \mathcal{K}(x^*, x)$ and $\mathrm{tr}(\cdot)$ is the trace of a matrix. Then, $\mathrm{IG}(x^*; X_m) \leq 0.5\log(v_*/(v_* - M_*))$.

By focusing on the simpler case of singleton $\mathcal{T}$, Proposition 1 provides an interpretable expression between the IG upper bound, the budget m and the observational constraint δ. In this IG upper bound, v_* is fixed given x^* and the kernel $\mathcal{K}$. Then, for a fixed budget m, a more restrictive observational constraint (i.e., a larger δ) leads to a smaller $\varepsilon_{\mathcal{K}}$ and M_*, resulting in a lower IG upper bound. On the other hand, for a fixed δ, a smaller budget m results in a lower IG upper bound (if the expression in the square brackets of M_* changes slowly/sublinearly with m). Put differently, the learning effectiveness of an agent without collaboration depends on the budget and its ability to collect data close to the target.

For an intuition of the proof, let $\widetilde{d}_* \coloneqq \det(\Sigma_{\mathcal{T}|X_m})/\det(K_{\mathcal{T}})$. Then,

$$\mathrm{IG}(\mathcal{T}; X_m) = -1/2\log\widetilde{d}_* \,. \tag{3}$$

Intuitively, the covariance matrix $K_{\mathcal{T}}$ captures the prior uncertainty at $\mathcal{T}$ whilst the covariance matrix $\Sigma_{\mathcal{T}|X_m}$, conditioned on the acquired data X_m, captures the posterior uncertainty. The ratio $\widetilde{d}_*$ naturally arises as an indicator of learning effectiveness where a smaller ratio implies more effective learning (i.e., higher IG). Importantly, Eq. (3) sheds light on the analysis of IG where a larger δ ($\mathcal{S}$ cannot cover $\mathcal{T}$) leads to a higher $\det(\Sigma_{\mathcal{T}|X_m})$ and thus a higher $\widetilde{d}_*$, resulting in a smaller (upper bound of) $\mathrm{IG}(\mathcal{T}; X_m)$. Note that this intuition applies to general $\mathcal{T}$ (i.e., not only singleton). However, the non-single $\mathcal{T}$ case is fundamentally more complex (due to additional cross-term covariances) and has a less interpretable IG upper bound due to the use of additional proof techniques (e.g., Lemma 4 in Appendix A).

Proposition 2. For general $\mathcal{T} = \{x^*_1, \ldots, x^*_{m'}\}$, budget m and $\inf_{x\in\mathcal{S},x^*\in\mathcal{T}} \|x - x^*\|_2 \geq \delta$. Define $\varepsilon_1 = \varepsilon_1(\delta, \mathcal{K}) \coloneqq \sup_{x\in\mathcal{S},x^*\in\mathcal{T}} \mathcal{K}(x, x^*)$, $\varepsilon_2 \coloneqq \max_{p,q} \widetilde{K}^{-1}_{X_m}[p, q]$, $\xi_{\text{off}} \coloneqq \max_{p\neq q} |(AB)[p, q]|$, and $\xi_{\text{diag}} \coloneqq \max_p |(AB)[p, p]|$ where $A \coloneqq K_{X_m\mathcal{T}}K^{-1}_{\mathcal{T}}K_{\mathcal{T}X_m}$, and $B \coloneqq \widetilde{K}^{-1}$. If $\max(\xi_{\text{diag}}, \xi_{\text{off}}) \leq \lambda_{\min}[K_{\mathcal{T}}]mm'\varepsilon_1^2\varepsilon_2$ where $\lambda_{\min}[\cdot]$ denotes the minimum eigenvalue of a matrix. Then,

$$\mathrm{IG}(\mathcal{T}; X_m) \leq 0.5(m-1)\times \\ \log[(1 - \xi_{\text{diag}} - (m-1)\xi_{\text{off}})(1 - \xi_{\text{diag}} + \xi_{\text{off}})]^{-1} \,.$$

In Proposition 2, the IG upper bound increases linearly with m but decreases at a logarithmic rate w.r.t. m, so in general, a larger m leads to a higher upper bound. However, The dependence on the observational constraint δ is more complicated: While ε_1 is monotonically decreasing in δ, ε_1 does not directly appear in the upper bound but serves in a condition. Assuming the condition on $\xi_{\text{diag}}, \xi_{\text{off}}$ holds, a smaller ε_1 implies both $\xi_{\text{diag}}, \xi_{\text{off}}$ are smaller. While a smaller ξ_{diag} does lead to a lower IG upper bound, the effect of a smaller ξ_{off} is worth future exploration. As we focus on how CAL can benefit the agents instead of the specific mechanics of single agent AL (e.g., how its observational constraint affects its IG), we adopt this IG upper bound thereafter, and defer further investigation to future work.

3.1.2 Collaboration Overcomes these Limitations

Denote the IG upper bound from Proposition 2 as $\overline{\mathrm{IG}}_{i,\text{indiv}}$. As mentioned previously, Eq. (2) enables the agents to help each other if their supports and targets complement each other to achieve a higher IG, formalised via a lower bound for $\mathrm{IG}_{i,m}$. Moreover, we derive the conditions where the lower bound for $\mathrm{IG}_{i,m} \geq \overline{\mathrm{IG}}_{i,\text{indiv}}$ to show IR.

Lemma 1. For agent i's target $\mathcal{T}_i = \{x^*_1, \ldots, x^*_{m'}\}$, and $X \coloneqq \{x_1, \ldots, x_m\}$ as some acquired data. Let $\bar{\varepsilon} \coloneqq \max_{x_l\in X, x^*_{l'}\in\mathcal{T}_i} \sup_{x\in\mathcal{X}} |\mathcal{K}(x_l, x) - \mathcal{K}(x^*_{l'}, x)|$, and $\widetilde{K}_X$ be defined w.r.t. X as in Sec. 2, then,

$$\mathrm{IG}(\mathcal{T}_i; X) \geq 1/2[-m\log(2\bar{\varepsilon} + m\bar{\varepsilon}^2 - \lambda) + \log\det(\widetilde{K}_X)] \,.^3$$

We highlight that X differs from X_m as X is a hypothetical data set for analysis purposes, so it does not have the same

[3] Note that since $\bar{\varepsilon}$ and λ are believed to be small, it is expected $2\bar{\varepsilon} + m\bar{\varepsilon}^2 - \lambda \leq 1$ (though not explicitly guaranteed).

observational constraint. Lemma 1 characterises a lower bound of IG w.r.t. $\mathcal{T}_i$ via X. Note that $\widetilde{K}_X, \lambda$ do not depend on $\mathcal{T}_i$, so whether X is useful to learn about $\mathcal{T}_i$ depends on $\bar{\varepsilon}$ and m. In particular, as $\bar{\varepsilon}$ relates the target to the acquired data through their kernel values w.r.t. an arbitrary $x \in \mathcal{X}$, a lower $\bar{\varepsilon}$ means X is useful to learn about $\mathcal{T}_i$ (i.e., a higher IG lower bound). In addition, a larger m (more data) gives a higher IG lower bound.

An important implication of Lemma 1 is that if it is possible in a collaboration to acquire data similar to X s.t. the IG lower bound is higher than $\overline{\text{IG}}_{i,\text{indiv}}$, then the collaboration satisfies IR as agent i has a higher IG than the IG upper bound without collaboration.

Proposition 3. For $i \in N$, if there is a covering set $X_{\text{cover},i} \subseteq X_{N,m}$ for $\mathcal{T}_i$, then $\overline{\text{IG}}_{i,\text{indiv}} \leq \text{IG}(\mathcal{T}_i, X_{N,m})$.

The full technical definition of a covering set $X_{\text{cover},i}$ for $\mathcal{T}_i$ is deferred to Appendix A (Definition 2) because it directly combines Proposition 2 and Lemma 1 to describe a set of data that "covers" $\mathcal{T}_i$ so that $\text{IG}(\mathcal{T}_i; X_{\text{cover},i}) \geq \overline{\text{IG}}_{i,\text{indiv}}$. Then, Proposition 3 states that if collaboration enables the agents to acquire data such that each target $\mathcal{T}_i$ can be covered, then agent i has a higher IG in collaboration than without collaboration. To see this, by definition of the covering set, $\text{IG}(\mathcal{T}_i; X_{\text{cover},i}) \geq \overline{\text{IG}}_{i,\text{indiv}}$; since Lemma 1 also implies that more data leads to a higher IG lower bound and $X_{\text{cover},i} \subseteq X_{N,m}$, the result follows. To illustrate, suppose the supports of the agents are largely different (even completely disjoint), but importantly these supports "cover" the targets of some other agents, then the collaboration will be effective (empirically shown in Sec. 4). In other words, our method does *not* require the supports of the agents to be close or similar.

However, we highlight that the condition that a covering set exists for each agent i is difficult to explicitly guarantee during the optimisation in Eq. (2). This difficulty is because the optimisation objective is specifically designed to provide a near-optimal performance (Lemma 6 in Appendix A). This implies a modification to the optimisation to explicitly guarantee the condition may "break" the near-optimality. Nevertheless, we empirically observe that the IGs of the agents improve with collaboration (e.g., Table 1 in Sec. 4), which verifies the effectiveness of Eq. (2) and Proposition 3.

3.2 Fairness via Maximising Nash Social Welfare

As in Sec. 1, fairness means all agents have equitable IGs, which (seemingly) requires equitable rates of learning for all the agents in every iteration. Then, by leveraging Eq. (2)'s equivalence to maximising a (modified) Nash social welfare (NSW), we show that Eq. (2) helps achieve equitable rates of learning for all the agents in an iteration. Finally, we adopt a global-to-local paradigm to show Eq. (2) helps achieve equitable final IGs, thus fairness.

From Eq. (2), the rate of learning for agent i in iteration k is the marginal IG $\Delta_{i,k}$. A more convenient (but equivalent) form is derived from $\widetilde{d}^{-1}_{*,i,k} = \exp(2\text{IG}_{i,k})$ from Eq. (3) with explicit notational dependence on i, k. Let $\rho_{i,k} := \widetilde{d}^{-1}_{*,i,k} / \widetilde{d}^{-1}_{*,i,k-1}$, then $\log \rho_{i,k} = \Delta_{i,k}$. Since the IGs before the collaboration are all 0, then intuitively, if the rates of learning (or marginal IGs $\Delta_{i,k}$) are equitable in each iteration, then the final IGs are also equitable:

Claim 1. Let $\delta_{\text{IG}} \geq 0$, if for all $1 \leq k \leq m$, $m \max(\frac{\rho_{i,k}}{\rho_{j,k}}, \frac{\rho_{j,k}}{\rho_{i,k}}) \leq \exp(\delta_{\text{IG}})$, then $|\text{IG}_{i,m} - \text{IG}_{j,m}| \leq \delta_{\text{IG}}$.

Therefore, fairness can be achieved if the rates of learning $\{\rho_{i,k}\}_{i \in N}$ are equitable. To see why $\rho_{i,k}$ is more convenient in showing Eq. (2) achieves fairness, re-write Eq. (2):

$$\text{argmax} \sum_i \log(\rho_{i,k})/\beta_i = \text{argmax} \prod_i \rho_{i,k}^{1/\beta_i} , \quad (4)$$

which corresponds to maximising the $\{1/\beta_i\}$-powered NSW of $\rho_{i,k}$'s. If all β_i's are equal, it reduces to NSW (Kaneko and Nakamura, 1979), which is often adopted to jointly optimise welfare and fairness; the welfare in this case is formalised via the sum of marginal IGs, and the fairness is in terms of how equitable the $\rho_{i,k}$'s are. Informally, NSW favours high total rates of learning but simultaneously encourages inequality-reducing transfers (Endriss, 2018). As an example with two agents (fix a k), while a utilitarian approach may prefer $\boldsymbol{\varrho}_k = (1.201, 1)$ to $\boldsymbol{\rho}_k = (1.1, 1.1)$ as $1.201 + 1 > 1.1 + 1.1$, NSW prefers $\boldsymbol{\rho}_k$ because $1.1 \times 1.1 > 1.201 \times 1$. Importantly, (maximising) NSW helps ensure that $\rho_{i,k}$'s are equitable as it evaluates the product instead of the sum (de Clippel, 2010), formalised by the Pigou Dalton Principle (PDP) (Sakamoto, 2020). Formally, Eq. (2) satisfies a generalised PDP.

Definition 1 (ε**-Pigou Dalton Principle).** A social welfare ordering (SWO) ι satisfies ε-PDP if for any $\boldsymbol{\rho} := \{\rho_i; i \in N\}$ and $\boldsymbol{\varrho} := \{\varrho_i; i \in N\}$ (the subscript k is omitted): $\forall \varepsilon' > \varepsilon \geq 0, (\exists i, j \in N, \varrho_i - \varepsilon' = \rho_i \geq \rho_j = \varrho_j + \varepsilon') \wedge (\forall l \in N \setminus \{i, j\}, \rho_l = \varrho_l) \implies \boldsymbol{\rho} \succeq_\iota \boldsymbol{\varrho}$.

This definition formalises the example of two agents and $\boldsymbol{\rho} \succeq_\iota \boldsymbol{\varrho}$ means ι (e.g., NSW) prefers $\boldsymbol{\rho}$ to $\boldsymbol{\varrho}$. As $\varepsilon \to 0$, ε-PDP recovers PDP, so a smaller ε is more desirable.

Lemma 2. Define $\zeta_{\bar{\beta}} := \max_{i,j \in N: 1/\beta_i > 1/\beta_j} (\beta_i - \beta_j)/\beta_i$ and $\varepsilon_{i,j} := \inf\{\varepsilon'; \varepsilon' > 0 \wedge \boldsymbol{\rho} \in [1, \infty)^n \wedge \rho_i \geq \rho_j \wedge (\rho_i - \varepsilon')(\rho_j + \varepsilon') \geq (\rho_i \rho_j)^{1+\zeta_{\bar{\beta}}}\}$. Then Eq. (2) satisfies ε_β-PDP with $\varepsilon_\beta = \max_{i,j \in N: i \neq j} \varepsilon_{i,j}$.

Lemma 2 characterises the effects of the sharing coefficients β_i on fairness in a single iteration (since the subscript k is omitted). Recall a smaller ε_β is more desirable. The smallest $\varepsilon_\beta = 0$ is when $\forall i, j, \beta_i = \beta_j$ and (2) reduces to NSW. Intuitively, it means if all agents are equally willing to share, then it is easier to achieve fairness. In contrast, if the differences in sharing coefficients are large (relative to β_i) then the fairness is weaker (ε_β is larger). In the extreme

case, suppose $1/\beta_i \to 1, 1/\beta_j \to 0,\ \forall j \neq i$, then IG for agent i is optimised with highest priority while IGs for other agents are not explicitly optimised. Empirically, our method is relatively robust to the differences in β_i's (equitable IGs even $\beta_i = 10\beta_j$ for some $i \neq j$, e.g., Table 1 in Sec. 4).

Global-to-local paradigm. Recall Claim 1 requires equitable $\{\rho_{i,k}\}_{i\in N}$ "globally" over all iterations $1 \leq k \leq m$, but Lemma 2 describes a fairness "localised" to the k-th iteration. To globalise the localised fairness to all iterations, define an SWO as follows,

$$\iota_{\text{global}}(\boldsymbol{\rho}_m) \coloneqq \textstyle\prod_{i\in N}\prod_{k=1}^{m}\rho_{i,k} \tag{5}$$

where $\boldsymbol{\rho}_m = \rho_{1,1},\dots,\rho_{n,1},\dots,\rho_{1,m},\dots,\rho_{n,m}$ and ι_{global}, $\boldsymbol{\rho}_m \succeq_{\iota_{\text{global}}} \boldsymbol{\varrho}_m$ iff $\prod_{i\in N}\prod_{k=1}^{m}\rho_{i,k} \geq \prod_{i\in N}\prod_{k=1}^{m}\varrho_{i,k}$. It extends Eq. (4) with the additional product over the m iterations. Consequently, as Eq. (4) helps achieve fairness in iteration k (Lemma 2), Eq. (5) is used to show a fairness guarantee over m iterations.

Proposition 4. $\iota_{\text{global}}(\boldsymbol{\rho}_m)$ in Eq. (5) satisfies $\varepsilon_{\text{global}}$-PDP where $\varepsilon_{\text{global}} \coloneqq \max_{1\leq k\leq m}\varepsilon_{\beta,k}$ and $\varepsilon_{\beta,k}$ is ε_β in Lemma 2 but with an explicit notational dependency on iteration k.

We make two observations. Firstly, the condition (equitable rates of learning *every* iteration k) in Claim 1 can be relaxed to equitable rates of learning in *aggregate* (i.e., the product over iterations). Moreover, we highlight that optimising an SWO that satisfies the generalised PDP naturally favours more equitable outcomes/IGs (such as the previous example with two agents) instead of definitively guaranteeing equitable outcomes/IGs. Empirically, our approach, in general, achieves the most equitable final IGs (e.g., Table 3 in Sec. 4).

Secondly, $\varepsilon_{\text{global}} = \max_{1\leq k\leq m}\varepsilon_{\beta,k}$ hints at a possible limitation: If in some iteration k, the fairness guarantee is poor (i.e., a large $\varepsilon_{\beta,k}$) due to the difficulties in optimising Eq. (2) (e.g., the high computational complexity for exact optimisation), it results in poor global fairness (i.e., high $\varepsilon_{\text{global}}$). Similar to the previous discussion of guaranteeing the condition of a covering set in Proposition 3, addressing this limitation (e.g., by modifying Eq. (2)) may "break" the near-optimality of $X_{N,m}$ and is thus challenging. Empirically, we find that using a heuristic to vary β_i over iterations can address this limitation, but its learning performance decreases slightly. Hence, how to attain both is an interesting future direction.

4 EMPIRICAL RESULTS

We empirically demonstrate that our algorithm outperforms existing baselines w.r.t. learning performance (via IR and subsequently regression error) and fairness on synthetic data and three real-world scientific discovery datasets. Our implementation can be found at https://github.com/XinyiYS/FAIR.

4.1 Experiment Setup

Datasets. We begin with synthetic data (1- and 2-dimensional) in order to specify the supports and targets and clearly interpret the outcome via visual illustrations. Subsequently, we consider several real-world scientific discovery datasets, including material design (MD) (Bassman et al., 2018), drug discovery (DD) (Pahikkala et al., 2015), and differential equation discovery (DED) (Heinonen et al., 2018). We consider $n = 3$ agents to be able to quantitatively and qualitatively verify the results without losing generality (Sim et al., 2020; Xu et al., 2021b).

MD involves predicting the real-valued *band gap* of a synthesised 3-layer material of different atoms. However, the agents are unable to synthesise certain materials due to the lack of access to certain atoms (i.e., the observational constraints) and, thus, unable to learn to predict their corresponding band gaps. DD involves predicting the real-valued *affinity score* between certain drug molecules and amino acids. Similarly, the agents are unable to synthesise certain drugs (i.e., the observational constraints) and thus unable to learn to predict their corresponding affinity scores. DED requires the agents to learn an underlying (possibly stochastic) 2-dimensional differential equation (DE) from the data. However, the agents are unable to observe the full trajectory, which can span from time $t = 0$ to $t = 10$, but an agent can only observe within $t \in [0, 2]$. Further, there are 3 separate DE systems denoted by ODE, VDP, and SDE where the true underlying DE is deterministic in ODE but stochastic in both VDP and SDE (i.e., more difficult to learn). Additional details on the GP implementation are deferred to Appendix B.

These real-world datasets echo the use cases in Sec. 1 where the agents (researchers) can benefit through collaboratively performing experiments and sharing the results because the observational constraints prevent them from learning effectively (e.g., cannot directly synthesise certain drugs) or cost-effectively (e.g., expensive to perform a large number of experiments) by themselves.

Baselines. We consider the following baselines: a batch AL method (Chen and Krause, 2013) (joint), entropy maximising acquisition function (Lewis and Catlett, 1994) (entropy), random acquisition function (rand), and individual without collaboration for verifying IR (ind). For our method via Eq. (2), we vary the sharing coefficients $[10, 5, 1], [10, 2, 1], [10, 1, 1], [1, 1, 1]$ (denoted as greedy_1-4, respectively) for the synthetic data. The coefficients for real-world datasets are in Appendix B. Moreover, inspired by Proposition 4, we include a heuristic that dynamically changes β_i over the iterations k: $\beta_{i,k+1} \leftarrow 0.5[\beta_{i,k} + \exp(\text{IG}_{i,k})/(\sum_i \exp(\text{IG}_{i,k}))]$, to investigate the IG vs. fairness trade-off (denoted as dynamic β). Intuitively, it sets a *higher* sharing coefficient $\beta_{i,k+1}$ for i if $\text{IG}_{i,k}$ is already relatively *high* (so i may be more willing to share

and let other agents improve their IGs).

Implementation of Eq. (2). The *exact* optimisation over $\mathcal{S}_N$ in Eq. (2) can be practically intractable for large n and/or $\mathcal{S}_i$. Therefore, our implementation adopts an approximation by first drawing a random subset $\mathcal{S}'_N \subset \mathcal{S}_N$ and optimising over $\mathcal{S}'_N$: $\text{argmax}_{\vec{x}_{N,k} \in \mathcal{S}'_N}$ in Eq. (2). The size of the subset $\mathcal{S}'_N$ is 1000 for synthetic datasets, DD and MD, and reduced to 10 for DED for computational reasons. Details (including a theoretical guarantee) are deferred to Appendix B.

Evaluation. For **learning performance**, (1) the final average IG as $\overline{\text{IG}}_m := (1/n)\sum_i \text{IG}_{i,m}$; (2) on the real-world datasets, the average (over the agents) of mean squared errors (MSEs) on $\mathcal{T}_i$: $\overline{\text{MSE}} := (1/n)\sum_i \text{MSE}(\mathcal{T}_i)$ where $\text{MSE}(\mathcal{T}_i)$ is the prediction error for agent i by performing GP regression on the acquired data and represents the final learning performance for i. For IG (MSE), higher (lower) is better. It is used to verify IR and provide assurance to the agents of performance improvement with collaboration. For **fairness**, (1) the mean-corrected standard deviation (std) of the final IGs: $\text{std}_{\text{IG}} := \text{std}(\{\text{IG}_{i,m}/\overline{\text{IG}}_m\}_i)$ (Li et al., 2020),[4] and (2) the std of the individual MSEs: $\text{std}_{\text{MSE}} := \text{std}(\{\text{MSE}(\mathcal{T}_i)\}_i)$. For $\text{std}_{\text{IG}}, \text{std}_{\text{MSE}}$, lower is better. We report the average and standard error (SE) over 5 independent trials. Verifying fairness (i.e., equitable final performance) is important to ensure that the agents would not be exploited (e.g., sharing data with others without getting an equitable performance).

4.2 Results

Synthetic data. We examine two settings for synthetic data (2D) called mismatch and identical. The mismatch setting considers the mismatch between each agent's $\mathcal{T}_i, \mathcal{S}_i$ (but with a match $\mathcal{T}_i, \mathcal{S}_j$ for $i \neq j$) to verify the quid-pro-quo benefit in Sec. 1 and the example shown in Fig. 1. The outcome from the collaboration is shown in Fig. 2. In both types of collaboration, the agents "share" their information to improve the learning effectiveness: Their jointly acquired observations can cover their respective targets significantly better than without collaboration as in Fig. 1 (right). The identical setting is when all agents have the (approximately) same $\mathcal{T}_i$, and if they have *sufficient* budget, each i can learn about $\mathcal{T}_i$ reasonably well. However, the agents have limited budgets by themselves (in this experiment $m = 5$), so they can benefit from sharing data to save the high cost of performing more experiments to collect data individually (the cost-saving benefit in Sec. 1).

Table 1 shows that our method performs best w.r.t. learning (i.e., high $\overline{\text{IG}}_m$) and fairness (i.e., low std_{IG}). In general, the high learning performance of our method (i.e., greedy_1-4) can be attributed to the near-optimality guarantee (Lemma 6 in Appendix A). Though the joint baseline achieves comparable performance in the mismatch setting, it underperforms our method in the identical setting and subsequently on real data.

[4]The fairness metric is not included for the ind baseline.

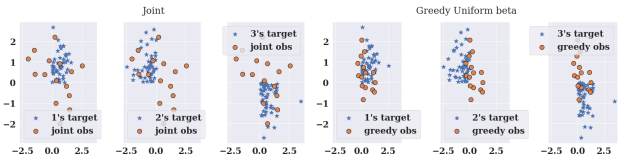

Figure 2: In the same setup as Fig. 1. Left: The agents collaborate by optimizing jointly (Chen and Krause, 2013). Right: Our method following Eq. (2).

From Proposition 4, one might expect that greedy_4 (i.e., $\forall i, j, \beta_i = \beta_j$) achieves the best fairness, but it is not the case for the mismatch setting. The reason is that the targets $\mathcal{T}_i \neq \mathcal{T}_j$, which can imply the inherent difficulties to learn about the different targets. Hence, having the same sharing coefficients as in greedy_4 might not guarantee the best fairness. Instead, a more careful way to update the sharing coefficients according to this difference (to account for the difficulties of learning about different $\mathcal{T}_i$'s) can lead to better fairness by dynamic β. In contrast, in the identical setting where $\mathcal{T}_i$ and $\mathcal{T}_j$ are very similar, setting the same sharing coefficients does achieve the best fairness.

Table 1: Learning performance (IR) and fairness (std) for **Synthetic-2D** mismatch (left) and identical (right). Average (SE) over 5 independent trials is reported (in brackets).

Setting	mismatch		identical	
Baselines	$\overline{\text{IG}}_m$	std_{IG}	$\overline{\text{IG}}_m$	std_{IG}
ind	42.130 (3.639)	N.A.	20.537 (4.777)	N.A.
rand	123.407 (2.739)	0.031 (0.022)	68.963 (0.413)	0.008 (0.006)
entropy	99.576 (2.314)	0.033 (0.023)	32.536 (1.088)	0.047 (0.033)
joint	143.570 (3.353)	0.033 (0.023)	71.093 (0.621)	0.012 (0.009)
greedy_1	133.613 (6.105)	0.065 (0.046)	81.847 (3.171)	0.055 (0.039)
greedy_2	138.877 (1.799)	0.018 (0.013)	81.884 (3.035)	0.052 (0.037)
greedy_3	140.196 (2.312)	0.023 (0.016)	82.423 (1.371)	0.024 (0.017)
greedy_4	**143.726 (4.202)**	0.041 (0.029)	**83.346 (0.402)**	**0.007 (0.005)**
dynamic β	139.623 (0.693)	**0.007 (0.005)**	81.876 (0.679)	0.012 (0.008)

Real data. For MD and DD, Tables 2 and 3 show that our method performs the best in terms of both learning performance ($\overline{\text{IG}}_m$ is higher by a considerable margin) and fairness (low std_{IG} and std_{MSE}). In addition, while dynamic β achieves the lowest std_{IG} for both MD and DD, it underperforms our method in terms of $\overline{\text{IG}}_m$ (and subsequently $\overline{\text{MSE}}$). This can be attributed to the near-optimality guarantee of Eq. (2) (Lemma 6 in Appendix A), which dynamic β does not provide.

On the most complicated data DED, Tables 4 and 5 show that collaboration (regardless of the baseline) results in a significant improvement since all baselines perform much better than the individual. Results for SDE are deferred to Appendix B. This implies that collaboration

Table 2: Learning performance (IG, MSE) and fairness evaluation (stds) for **MD**. Average (SE) over 5 independent trials is reported (in brackets).

Baselines	$\overline{\text{IG}}_m$	std_{IG}	$\overline{\text{MSE}}$	std_{MSE}
ind	51.708 (7.369)	N.A.	0.228 (0.001)	N.A.
rand	171.949 (7.276)	0.060 (0.042)	0.201 (0.018)	0.044 (0.004)
entropy	99.825 (2.176)	0.031 (0.022)	0.212 (0.016)	0.028 (0.006)
joint	182.095 (7.227)	0.056 (0.040)	0.209 (0.026)	0.040 (0.004)
greedy_1	215.673 (29.651)	0.194 (0.137)	0.222 (0.021)	0.031 (0.005)
greedy_2	220.073 (31.980)	0.206 (0.145)	0.262 (0.022)	0.054 (0.009)
greedy_3	221.808 (32.490)	0.207 (0.146)	0.226 (0.027)	**0.027 (0.007)**
greedy_4	**223.263 (14.591)**	0.092 (0.065)	**0.173 (0.024)**	0.035 (0.012)
dynamic β	215.206 (1.041)	**0.007 (0.005)**	0.255 (0.036)	0.046 (0.014)

Table 3: Learning performance (IG, MSE) and fairness evaluation (stds) for **DD**. Average (SE) over 5 independent trials is reported (in brackets).

Baselines	$\overline{\text{IG}}_m$	std_{IG}	$\overline{\text{MSE}}$	std_{MSE}
ind	1.577 (0.068)	N.A.	0.042 (0.000)	N.A.
rand	21.430 (6.594)	0.435 (0.308)	0.037 (0.005)	0.032 (0.007)
entropy	0.238 (0.068)	0.407 (0.288)	0.057 (0.000)	0.055 (0.000)
joint	56.872 (17.508)	0.435 (0.308)	0.026 (0.004)	0.020 (0.004)
greedy_1	65.174 (21.093)	0.458 (0.324)	0.033 (0.006)	0.029 (0.007)
greedy_2	65.617 (20.584)	0.444 (0.314)	**0.022 (0.004)**	0.022 (0.006)
greedy_3	65.207 (20.559)	0.446 (0.315)	0.024 (0.004)	0.018 (0.004)
greedy_4	**66.626 (22.223)**	0.472 (0.334)	**0.022 (0.002)**	**0.017 (0.004)**
dynamic β	54.261 (12.927)	**0.337 (0.238)**	0.034 (0.006)	0.028 (0.007)

is more beneficial for more challenging learning tasks and thus encourages collaboration among the agents (i.e., scientists/researchers). Our method, in particular, achieves the best learning performance in terms of $\overline{\text{IG}}_m$ and the best/competitive performance in terms of $\overline{\text{MSE}}$. The highest IG does not guarantee the lowest MSE because IG does *not* include the true regression labels, but MSE does.

Table 4: Learning performance (IG, MSE) and fairness evaluation (stds) for **DED-ODE**. Average (SE) over 5 independent trials is reported (in brackets).

Baselines	$\overline{\text{IG}}_m$	std_{IG}	$\overline{\text{MSE}}$	std_{MSE}
ind	-6.617 (0.393)	N.A.	7.270 (0.198)	N.A.
rand	2.011 (1.267)	0.891 (0.630)	4.506 (0.163)	2.342 (0.207)
entropy	0.832 (0.383)	0.651 (0.460)	**4.116 (0.132)**	2.524 (0.091)
joint	2.888 (1.057)	0.517 (0.366)	5.551 (0.803)	2.190 (0.465)
greedy_1	3.349 (0.484)	0.204 (0.144)	4.415 (0.070)	2.230 (0.238)
greedy_2	3.487 (0.480)	0.195 (0.138)	4.349 (0.238)	**2.136 (0.172)**
greedy_3	3.175 (0.374)	**0.167 (0.118)**	4.509 (0.189)	2.264 (0.224)
greedy_4	**3.706 (0.877)**	0.335 (0.237)	4.234 (0.032)	2.629 (0.053)
dynamic β	3.107 (0.551)	0.251 (0.177)	4.321 (0.103)	2.438 (0.130)

Conveniently, DED admits a visualisation in Fig. 3 to illustrate how and why collaboration is beneficial. Fig. 3 compares the learned trajectory w.r.t. the 1st dimension of the true DE with (left) and without collaboration (right). Intuitively, as the agent can only observe a partial trajectory (from a limited time window), it is unable to effectively learn the true DE (right). In contrast, through collaboration

Table 5: Learning performance (IG, MSE) and fairness evaluation (stds) for **DED-VDP**. Average (SE) over 5 independent trials is reported (in brackets).

Baselines	$\overline{\text{IG}}_m$	std_{IG}	$\overline{\text{MSE}}$	std_{MSE}
ind	-8.843 (0.050)	N.A.	6.796 (0.555)	N.A.
rand	2.855 (0.851)	0.432 (0.298)	4.974 (1.254)	2.457 (0.895)
entropy	2.775 (0.839)	0.428 (0.303)	4.473 (0.499)	1.860 (0.430)
joint	2.786 (0.872)	0.443 (0.313)	3.889 (0.299)	1.486 (0.345)
greedy_1	2.723 (0.845)	0.439 (0.310)	4.070 (0.369)	2.004 (0.511)
greedy_2	2.750 (0.822)	**0.423 (0.299)**	4.284 (0.140)	1.470 (0.448)
greedy_3	2.823 (0.877)	0.440 (0.311)	3.961 (0.367)	1.780 (0.357)
greedy_4	**2.899 (0.871)**	0.425 (0.300)	4.198 (0.320)	1.841 (0.355)
dynamic β	2.833 (0.855)	0.427 (0.302)	**3.831 (0.247)**	**1.335 (0.487)**

(i.e., sharing of data/observations over the full trajectory of the DE), the agent can effectively learn about the true DE (i.e., the sampled orange trajectories match the true VDP in blue). An illustration over the entire 2-dimensional vector field is deferred to the Appendix B.

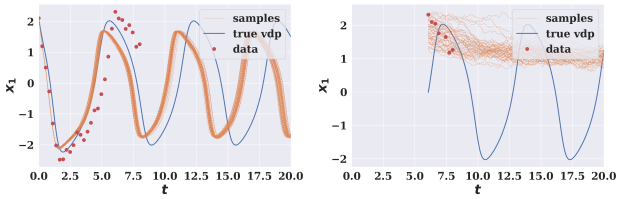

Figure 3: Left (right) shows sampled trajectories from the learnt GP of collaboration (agent 3 observing time $t \in [6, 8]$ w/o collaboration) w.r.t. the 1st dimension x_1 of the DE.

Key takeaways. Our empirical settings closely follow those motivated in Sec. 1 to reflect the observational constraints for agents due to (i) limited resources/budgets and/or (ii) no access to certain raw materials/geographical locations, etc. Our method is shown to enable effective collaborations by assuring the agents of better performance (higher $\overline{\text{IG}}_m$ and lower $\overline{\text{MSE}}$) than without collaboration for IR, and importantly, equitable performance (through low std_{IG} and std_{MSE}) for fairness. Empirically, dynamic β can provide better fairness by trading off some performance. It is interesting to investigate this further in the future.

5 RELATED WORK

Active learning. In addition to scientific discovery, the information-theoretic AL framework has also been applied to deep learning to leverage structural properties of neural networks (Gal et al., 2017; Ash et al., 2020). Subsequently, some existing works develop the batch version of AL (Chen and Krause, 2013) where instead of acquiring one data point per iteration, these methods acquire a batch of data points. However, typical batch AL methods have limited effectiveness in guaranteeing IR or fairness (shown in Sec. 4). Some other batch AL

methods (Kirsch et al., 2019; Gal et al., 2017) specifically target the classification problem, whereas we adopt the GP to leverage the closed-form expressions for IGs for regression. Separately, (Ash et al., 2020; Citovsky et al., 2021) have a different setting with homogeneous supports (our is heterogeneous) and a given parametric model, often a neural network (we do not assume such as model). Hence, their method is not directly comparable with ours.

Collaborative learning and data sharing. There have been an increasing number of recent works on the topic of collaborative (machine) learning. Sim et al. (2020) adopt a Bayesian perspective and facilitate collaborative learning by evaluating the reduction in entropy from including an agent's data. Hoang et al. (2021) propose a model-fusion method to enable collaborative learning from agents, each with a trained machine learning model (but without data). Xu et al. (2021a); Li et al. (2020, 2021) investigate the federated learning setting as a collaborative learning framework. Tay et al. (2022) study the unsupervised learning setting and propose a mechanism to incentivise the agents to share data/collaborate. Agussurja et al. (2022) consider the Bayesian inference framework and design an (asymptotically) Shapley-fair 2-agent data sharing/collaborative learning algorithm. In contrast, the motivating use case for our work (i.e., scientific discovery) and the setting (i.e., active learning in the information-theoretic framework) have not previously been explored in existing works.

In this line of work on collaborative learning, a notion of fairness is often emphasised. Sim et al. (2020); Xu et al. (2021a); Tay et al. (2022); Agussurja et al. (2022); Nguyen et al. (2022) all adopt the formalization of fairness from the well-known *Shapley value* (Shapley, 1953) from cooperative game theory. Intuitively, this notion of fairness encourages a sense of proportionality between what each agent shares/contributes and what each agent receives. In contrast, Sim et al. (2021); Li et al. (2020, 2021) champion an egalitarian notion of fairness by (approximately) equalising what each agent receives to reduce inequality and thus achieve fairness. Our work also shares this inequality-reducing perspective as the notion of fairness by exploiting a problem-specific structure that enables the maximisation of Nash social welfare. The use of Nash social welfare in collaborative learning has not been previously investigated.

Another distinction from some existing works is that our method explicitly considers the welfare, in addition to fairness, which in this case refers to the learning performance of the agents after the collaboration. In doing so, our work additionally guarantees individual rationality (Sim et al., 2020; Tay et al., 2022). Intuitively, it means the agents can achieve better learning performance with collaboration than without.

6 DISCUSSION AND CONCLUSION

We propose *collaborative active learning* formalised by a coordinated acquisition function to enable the agents (researchers) to collaborate during data collection. By leveraging the Gaussian process and an information-theoretic approach, we characterise the conditions for *individual rationality* (i.e., the information gain of an agent improves in collaboration). By utilising a modified Nash social welfare to satisfy a generalised Pigou Dalton Principle, we show that our coordinated acquisition function helps achieve fairness (i.e., the information gains of the agents are equitable). We identify that the coordinated acquisition function, which is designed to achieve a near-optimal performance, can be somewhat restrictive (i.e., difficult to explicitly guarantee the covering set condition and/or address the fairness limitation). Nevertheless, our experimental results show that our approach outperforms existing baselines in achieving both individual rationality and fairness. For our future work, we plan to generalise our approach to cater to non-myopic active learning (Cao et al., 2013; Hoang et al., 2014; Ling et al., 2016; Low et al., 2009, 2008, 2011), level set estimation (Nguyen et al., 2021), and active learning of a multi-output GP model (Zhang et al., 2016), and improve its scalability with the use of distributed/decentralized (Chen et al., 2012, 2013a,b, 2015; Hoang et al., 2016, 2019; Low et al., 2015; Ouyang and Low, 2018), online/stochastic (Hoang et al., 2015, 2017; Low et al., 2014; Xu et al., 2014; Yu et al., 2019a), or deep (Yu et al., 2019b, 2021) sparse GP models to represent the belief of the unknown ground truth function efficiently.

Acknowledgements

This research/project is supported by the National Research Foundation Singapore and DSO National Laboratories under the AI Singapore Programme (AISG Award No: AISG2-RP-2020-018) and by A*STAR under its RIE2020 Advanced Manufacturing and Engineering (AME) Programmatic Funds (Award A20H6b0151). Xinyi Xu is also supported by the Institute for Infocomm Research of Agency for Science, Technology and Research (A*STAR). The authors would like to thank Rachael Hwee Ling Sim for the valuable discussion and her constructive feedback.

References

Lucas Agussurja, Xinyi Xu, and Bryan Kian Hsiang Low. On the convergence of the Shapley value in parametric Bayesian learning games. In *Proc. ICML*, pages 180–196, 2022.

Jordan T. Ash, Chicheng Zhang, Akshay Krishnamurthy, John Langford, and Alekh Agarwal. Deep Batch Active Learning by Diverse, Uncertain Gradient Lower Bounds. In *Proc. ICLR*, 2020.

Steven Atkinson, Waad Subber, Liping Wang, Genghis Khan, Philippe Hawi, and Roger Ghanem. Data-driven discovery of free-form governing differential equations. In *Proc. NeurIPS Workshop on Machine Learning and the Physical Sciences*, 2019.

Lindsay Bassman, Pankaj Rajak, Rajiv K. Kalia, Aiichiro Nakano, Fei Sha, Jifeng Sun, David J. Singh, Muratahan Aykol, Patrick Huck, Kristin Persson, and Priya Vashishta. Active learning for accelerated design of layered materials. *npj Computational Materials*, 4:74, 2018.

Olivier Borkowski, Mathilde Koch, Agnès Zettor, Amir Pandi, Angelo Cardoso Batista, Paul Soudier, and Jean Loup Faulon. Large scale active-learning-guided exploration for in vitro protein production optimization. *Nature Communications*, 11:1872, 2020.

Richard P. Brent, Judy Anne H. Osborn, and Warren D. Smith. Note on best possible bounds for determinants of matrices close to the identity matrix. *Linear Algebra and its Applications*, 466:21–26, 2015.

N. Cao, K. H. Low, and J. M. Dolan. Multi-robot informative path planning for active sensing of environmental phenomena: A tale of two algorithms. In *Proc. AAMAS*, 2013.

J. Chen, K. H. Low, C. K.-Y. Tan, A. Oran, P. Jaillet, J. M. Dolan, and G. S. Sukhatme. Decentralized data fusion and active sensing with mobile sensors for modeling and predicting spatiotemporal traffic phenomena. In *Proc. UAI*, pages 163–173, 2012.

J. Chen, N. Cao, K. H. Low, R. Ouyang, C. K.-Y. Tan, and P. Jaillet. Parallel Gaussian process regression with low-rank covariance matrix approximations. In *Proc. UAI*, pages 152–161, 2013a.

J. Chen, K. H. Low, and C. K.-Y. Tan. Gaussian process-based decentralized data fusion and active sensing for mobility-on-demand system. In *Proc. RSS*, 2013b.

J. Chen, K. H. Low, P. Jaillet, and Y. Yao. Gaussian process decentralized data fusion and active sensing for spatiotemporal traffic modeling and prediction in mobility-on-demand systems. *IEEE Trans. Autom. Sci. Eng.*, 12:901–921, 2015.

Yuxin Chen and Andreas Krause. Near-optimal batch mode active learning and adaptive submodular optimization. In *Proc. ICML*, pages 160–168, 2013.

Gui Citovsky, Giulia DeSalvo, Claudio Gentile, Lazaros Karydas, Anand Rajagopalan, Afshin Rostamizadeh, and Sanjiv Kumar. Batch active learning at scale. In *Proc. NeurIPS*, pages 11933–11944, 2021.

Erik A. Daxberger and Bryan Kian Hsiang Low. Distributed batch Gaussian process optimization. In *Proc. ICML*, pages 951–960, 2017.

Geoffroy de Clippel. Comment on "The Veil of Public Ignorance". Working papers 2010-3, Department of Economics, Brown University, 2010.

Brian M. de Silva, David M. Higdon, Steven L. Brunton, and J. Nathan Kutz. Discovery of physics from data: Universal laws and discrepancies. *Frontiers in Artificial Intelligence*, 3, 2020.

Marcia Dunn. Moon goes blood red this weekend: 'Eclipse for the Americas', 2022. URL https://phys.org/news/2022-05-moon-blood-red-weekend-eclipse.html.

Frank Emmert-Streib, Matthias Dehmer, and Olli Yli-Harja. Against dataism and for data sharing of big biomedical and clinical data with research parasites. *Frontiers in Genetics*, 7, 2016.

Ulle Endriss. Lecture notes on fair division. arXiv:1806.04234, 2018.

Yarin Gal, Riashat Islam, and Zoubin Ghahramani. Deep Bayesian active learning with image data. In *Proc. ICML*, pages 1923–1932, 2017.

Konstantin Gubaev, Evgeny V. Podryabinkin, Gus L.W. Hart, and Alexander V. Shapeev. Accelerating high-throughput searches for new alloys with active learning of interatomic potentials. *Computational Materials Science*, 156:148–156, 2019.

Markus Heinonen, Cagatay Yildiz, Henrik Mannerström, Jukka Intosalmi, and Harri Lähdesmäki. Learning unknown ODE models with Gaussian processes. In *Proc. ICML*, pages 1959–1968, 2018.

Q. M. Hoang, T. N. Hoang, and K. H. Low. A generalized stochastic variational Bayesian hyperparameter learning framework for sparse spectrum Gaussian process regression. In *Proc. AAAI*, pages 2007–2014, 2017.

T. N. Hoang, K. H. Low, P. Jaillet, and M. Kankanhalli. Nonmyopic ε-Bayes-optimal active learning of Gaussian processes. In *Proc. ICML*, pages 739–747, 2014.

T. N. Hoang, Q. M. Hoang, and K. H. Low. A unifying framework of anytime sparse Gaussian process regression models with stochastic variational inference for big data. In *Proc. ICML*, pages 569–578, 2015.

T. N. Hoang, Q. M. Hoang, and K. H. Low. A distributed variational inference framework for unifying parallel sparse Gaussian process regression models. In *Proc. ICML*, pages 382–391, 2016.

T. N. Hoang, Q. M. Hoang, K. H. Low, and J. P. How. Collective online learning of Gaussian processes in massive multi-agent systems. In *Proc. AAAI*, pages 7850–7857, 2019.

Trong Nghia Hoang, Shenda Hong, Cao Xiao, Bryan Low, and Jimeng Sun. AID: Active distillation machine to

leverage pre-trained black-box models in private data settings. In *Proc. WWW*, page 3569–3581, 2021.

Kevin Maik Jablonka, Giriprasad Melpatti Jothiappan, Shefang Wang, Berend Smit, and Brian Yoo. Bias free multiobjective active learning for materials design and discovery. *Nature Communications*, 12:2312, 2021.

Mohammad Neamul Kabir and Limsoon Wong. EnsembleFam: towards more accurate protein family prediction in the twilight zone. *BMC Bioinformatics*, 23 (1):1–20, 2022.

Mamoru Kaneko and Kenjiro Nakamura. The nash social welfare function. *Econometrica*, 47(2):423–435, 1979.

Douglas B. Kell. Scientific discovery as a combinatorial optimisation problem: How best to navigate the landscape of possible experiments? *BioEssays*, 34(3):236–244, 2012.

Andreas Kirsch, Joost van Amersfoort, and Yarin Gal. BatchBALD: Efficient and diverse batch acquisition for deep Bayesian active learning. In *Proc. NeurIPS*, volume 32, 2019.

Andreas Krause and Carlos Guestrin. Near-optimal nonmyopic value of information in graphical models. In *Proc. UAI*, page 324–331, 2005.

Andreas Krause, Ajit Singh, and Carlos Guestrin. Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies. *JMLR*, 9:235–284, 2008.

Ravinder Kumar. *Bounds for eigenvalues*. PhD thesis, University of Alberta, 1984.

A. Gilad Kusne, Heshan Yu, Changming Wu, Huairuo Zhang, Jason Hattrick-Simpers, Brian DeCost, Suchismita Sarker, Corey Oses, Cormac Toher, Stefano Curtarolo, Albert V. Davydov, Ritesh Agarwal, Leonid A. Bendersky, Mo Li, Apurva Mehta, and Ichiro Takeuchi. On-the-fly closed-loop materials discovery via Bayesian active learning. *Nature Communications*, 11:5966, 2020.

David D. Lewis and Jason Catlett. Heterogeneous uncertainty sampling for supervised learning. In *Proc. ICML*, pages 148–156, 1994.

Tian Li, Maziar Sanjabi, Ahmad Beirami, and Virginia Smith. Fair resource allocation in federated learning. In *Proc. ICLR*, 2020.

Tian Li, Shengyuan Hu, Ahmad Beirami, and Virginia Smith. Ditto: Fair and robust federated learning through personalization. In *Proc. ICML*, pages 6357–6368, 2021.

C. K. Ling, K. H. Low, and P. Jaillet. Gaussian process planning with Lipschitz continuous reward functions: Towards unifying Bayesian optimization, active learning, and beyond. In *Proc. AAAI*, pages 1860–1866, 2016.

Bernard Lo and David L. DeMets. Incentives for clinical trialists to share data. *New England Journal of Medicine*, 375(12):1112–1115, 2016.

Dan L. Longo and Jeffrey M. Drazen. Data sharing. *New England Journal of Medicine*, 374(3):276–277, 2016.

Turab Lookman, Prasanna V. Balachandran, Dezhen Xue, and Ruihao Yuan. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Computational Materials*, 5:21, 2019.

K. H. Low, J. M. Dolan, and P. Khosla. Adaptive multi-robot wide-area exploration and mapping. In *Proc. AAMAS*, pages 23–30, 2008.

K. H. Low, J. M. Dolan, and P. Khosla. Information-theoretic approach to efficient adaptive path planning for mobile robotic environmental sensing. In *Proc. ICAPS*, 2009.

K. H. Low, J. M. Dolan, and P. Khosla. Active Markov information-theoretic path planning for robotic environmental sensing. In *Proc. AAMAS*, pages 753–760, 2011.

K. H. Low, N. Xu, J. Chen, K. K. Lim, and E. B. Özgül. Generalized online sparse Gaussian processes with application to persistent mobile robot localization. In *Proc. ECML/PKDD Nectar Track*, pages 499–503, 2014.

K. H. Low, J. Yu, J. Chen, and P. Jaillet. Parallel Gaussian process regression for big data: Low-rank representation meets Markov approximation. In *Proc. AAAI*, pages 2821–2827, 2015.

G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher. An analysis of approximations for maximizing submodular set functions – I. *Mathematical Programming*, 14: 265–294, 1978.

Q. P. Nguyen, B. K. H. Low, and P. Jaillet. An information-theoretic framework for unifying active learning problems. In *Proc. AAAI*, pages 9126–9134, 2021.

Quoc Phong Nguyen, Bryan Kian Hsiang Low, and Patrick Jaillet. Trade-off between payoff and model rewards in Shapley-fair collaborative machine learning. In *Proc. NeurIPS*, 2022.

R. Ouyang and K. H. Low. Gaussian process decentralized data fusion meets transfer learning in large-scale distributed cooperative perception. In *Proc. AAAI*, pages 3876–3883, 2018.

Tapio Pahikkala, Antti Airola, Sami Pietilä, Sushil Shakyawar, Agnieszka Szwajda, Jing Tang, and Tero Aittokallio. Toward more realistic drug-target interaction predictions. *Briefings in Bioinformatics*, 16(2):325–337, 2015.

YoSon Park and Casey S Greene. A parasite's perspective on data sharing. *GigaScience*, 7(11), 2018.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

David Rousseau and Andrey Ustyuzhanin. Machine learning scientific competitions and datasets. In Paolo Calafiura, editor, *Artificial Intelligence for High Energy Physics*, chapter 20, pages 765–812. 2022.

Norihito Sakamoto. Equity principles and interpersonal comparison of well-being: Old and new joint characterizations of generalized leximin, rank-dependent utilitarian, and leximin rules. RCNE Discussion Paper Series 7, Research Center for Normative Economics, Institute of Economic Research, Hitotsubashi University, 2020.

Gisbert Schneider. Automating drug discovery. *Nature Reviews Drug Discovery*, 17:97–113, 2018.

L. S. Shapley. A value for n-person games. In H. W. Kuhn and A. W. Tucker, editors, *Contributions to the Theory of Games*, volume 2, pages 307–317. Princeton Univ. Press, 1953.

Rachael Hwee Ling Sim, Yehong Zhang, Mun Choon Chan, and Bryan Kian Hsiang Low. Collaborative machine learning with incentive-aware model rewards. In *Proc. ICML*, pages 8927–8936, 2020.

Rachael Hwee Ling Sim, Yehong Zhang, Bryan Kian Hsiang Low, and Patrick Jaillet. Collaborative Bayesian optimization with fair regret. In *Proc. ICML*, pages 9691–9701, 2021.

Darren B. Taichman, Joyce Backus, Christopher Baethge, Howard Bauchner, Peter W. de Leeuw, Jeffrey M. Drazen, John Fletcher, Frank A. Frizelle, Trish Groves, Abraham Haileamlak, Astrid James, Christine Laine, Larry Peiperl, Anja Pinborg, Peush Sahni, and Sinan Wu. Sharing clinical trial data – a proposal from the international committee of medical journal editors. *New England Journal of Medicine*, 374:384–386, 2016.

Sebastian Shenghong Tay, Xinyi Xu, Chuan Sheng Foo, and Bryan Kian Hsiang Low. Incentivizing collaboration in machine learning via synthetic data rewards. In *Proc. AAAI*, pages 9448–9456, 2022.

Adam Voiland. The best places to see the eclipse, 2017. URL https://earthobservatory.nasa.gov/images/90729/the-best-places-to-see-the-eclipse.

N. Xu, K. H. Low, J. Chen, K. K. Lim, and E. B. Özgül. GP-Localize: Persistent mobile robot localization using online sparse Gaussian process observation model. In *Proc. AAAI*, pages 2585–2592, 2014.

Xinyi Xu, Lingjuan Lyu, Xingjun Ma, Chenglin Miao, Chuan Sheng Foo, and Kian Hsiang Low. Gradient driven rewards to guarantee fairness in collaborative machine learning. In *Proc. NeurIPS*, pages 16104–16117, 2021a.

Xinyi Xu, Zhaoxuan Wu, Chuan-Sheng Foo, and Bryan Kian Hsiang Low. Validation free and replication robust volume-based data valuation. In *Proc. NeurIPS*, pages 10837–10848, 2021b.

H. Yu, T. N. Hoang, K. H. Low, and P. Jaillet. Stochastic variational inference for Bayesian sparse Gaussian process regression. In *Proc. IJCNN*, 2019a.

H. Yu, D. Liu, K. H. Low, and P. Jaillet. Convolutional normalizing flows for deep Gaussian processes. In *Proc. IJCNN*, 2021.

Haibin Yu, Yizhou Chen, Zhongxiang Dai, Bryan Kian Hsiang Low, and Patrick Jaillet. Implicit posterior variational inference for deep Gaussian processes. In *Proc. NeurIPS*, pages 14475–14486, 2019b.

Y. Zhang, T. N. Hoang, K. H. Low, and M. Kankanhalli. Near-optimal active learning of multi-output Gaussian processes. In *Proc. AAAI*, pages 2351–2357, 2016.

A PROOFS AND DERIVATIONS

A.1 Useful Lemmas - Existing Results

Lemma 3 (Upper bound on matrix quadratic norm using eigenvalues). Let $A \in \mathbb{R}^{n\times n}$ and $A^\top = A$ and $\lambda_{\min}, \lambda_{\max}$ denote its maximum and minimum eigenvalues. Then,

$$\forall u \in \mathbb{R}^n, \lambda_{\min} u^\top u \leq u^\top A u \leq \lambda_{\max} u^\top u.$$

The upper- and lower bounds are tight because they can be realised by the corresponding eigenvectors to $\lambda_{\max}$ and $\lambda_{\min}$, respectively.

Proof. Since A is real and symmetric, it is orthogonally diagonalisable. Denote its orthonormal eigenvectors as $[\mu_1, \ldots, \mu_n]$ with corresponding eigenvalues as $[\lambda_1, \ldots, \lambda_n]$ (i.e., $A\mu_i = \lambda_i \mu_i$). Define matrix P such that its i-th column is μ_i and matrix D as $\text{diag}([\lambda_1, \ldots, \lambda_n])$ so that its i-th diagonal entry is λ_i. Then, we have $A = PDP^\top$ and $P^\top P = \mathbf{I}$.

For some $u \in \mathbb{R}^n$, let $v = P^\top u$, we have

$$v^\top v = (P^\top u)^\top P^\top u = u^\top P^\top P u = u^\top u$$

and

$$u^\top A u = u^\top P D P^\top u = (P^\top u)^\top D (P^\top u) = v^\top D v = \sum_{i=1}^{n} \lambda_i \mu_i^2.$$

W.l.o.g., assume $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n$, then

$$\sum_i \lambda_n \mu_i^2 \leq \sum_i \lambda_i \mu_i^2 \leq \sum_i \lambda_1 \mu_i^2 \implies \lambda_n v^\top v \leq \sum_i \lambda_i \mu_i^2 \leq \lambda_1 v^\top v$$

$$\implies \lambda_n u^\top u \leq u^\top A u \leq \lambda_1 u^\top u.$$

□

Lemma 4 (Bounds on the determinant of an approximate identity matrix (Brent et al., 2015, Theorem 1,2)). For a matrix $A = I - E \in \mathbb{R}^{t\times t}$ such that $|E_{i,j}| \leq \xi_{\text{off}}, i \neq j$ and $|E_{i,i}| \leq \xi_{\text{diag}}$, then

$$\det(A) \in [(1 - \xi_{\text{diag}} - (t-1)\xi_{\text{off}})(1 - \xi_{\text{diag}} + \xi_{\text{off}})^{t-1}, ((1 + \xi_{\text{diag}})^2 + (t-1)\xi_{\text{off}}^2)^{t/2}]$$

where the lower bound requires a condition $\xi_{\text{diag}} + \xi_{\text{off}}(t-1) \leq 1$ and the lower bound is sharp (best possible).

Lemma 5 (Matrix determinant lemma). For an invertible square matrix $A \in \mathbb{R}^{n\times n}$ and $U, V \in \mathbb{R}^{n\times m}$, we have

$$\det(A + UV^\top) = \det(\mathbf{I}_m + V^\top A^{-1} U) \det(A).$$

A.2 Useful Lemma - Our Result

Lemma 6 (Near-optimality of a monotonic submodular function via the greedy heuristic). The solution to Eq. (2) has $\sum_i \text{IG}_{i,m}/\beta_i \geq (1 - 1/e) \max_{X_{N,m}} \sum_i \text{IG}_{i,m}/\beta_i$.

Proof. The proof follows from (Krause et al., 2008, Theorem 7) by noting $\sum_i \text{IG}_{i,k}/\beta_i$ is monotone submodular because each individual $\text{IG}_{i,k}$ is monotone submodular $1/\beta_i > 0$. In general, a concave transformation of a monotone submodular function preserves monotonicity and submodularity. In this case, a non-negative linear combination using $1/\beta_i$ of submodular functions is a concave transformation.

Each individual $\text{IG}_{i,k}$ is monotone submodular because of the conditional independence assumption between (disjoint subsets of $\mathcal{S}_i$) and $\mathcal{T}$ by (Krause and Guestrin, 2005, Corollary 4). □

Note on the cumulative aspect of Lemma 6. Similar to (Krause et al., 2008, Theorem 7), Lemma 6 provides a cumulative performance guarantee in a sense for this greedy heuristic achieves a near-optimal factor of $1-1/e$ for each $1 \leq k \leq m$: $\sum_i \text{IG}_{i,k}/\beta_i \geq (1-1/e)\max_{X_{N,k}} \sum_i \text{IG}_{i,k}/\beta_i$. This cumulative guarantee is particularly useful in practice if (i) the outcome of each iteration k is important. For instance, each iteration takes a long time (e.g., due to long experimentation time), and after each iteration k, the collected data $X_{N,k}$ so far is used for analysis, learning another machine learning model for inference (Kabir and Wong, 2022), then the quality of $X_{N,k}$ is important for each iteration k, instead of only the final iteration; (ii) the budget m is not known *a priori*. For instance, a research project has a fixed timeline (e.g., to synthesise a vaccine/drug), but the time it takes to collect data (by performing experiments) is indeterminate beforehand, so the budget m might not be known. Then, it is important to ensure the quality of the collected data $X_{N,k}$ collected so far since it is not guaranteed there is sufficient time for the next iteration. Such advantages can further highlight the advantage of the cumulative aspect of this performance guarantee. For empirical verifications, we perform experiments to evaluate the cumulative performance in Appendix B (e.g., Fig. 6).

A.3 Proofs of Propositions 1 and 2

We first describe the high-level idea applicable to both singleton and general $\mathcal{T}$ and then present their respective derivations.

IG represents the reduction in uncertainty about $\mathcal{T}$ through some acquired data where, if the acquired data are close (i.e., by Euclidean distance) to $\mathcal{T}$, then IG is high. Consequently, if these data are far (constrained by the support, which is far from $\mathcal{T}$), then there is an upper bound (limit) on IG that depends on this distance between the support and $\mathcal{T}$ (which also depends on the kernel and its related hyperparameters).

Proof of Proposition 1. Define the kernel operation $k_* := \mathcal{K}(x^*, x^*)$, which is equal to the prior variance v^* on x^* by definition of GP. Let posterior variance on a single target x^* be $\sigma_m^2(x^*)$ (from the acquired data $X_m \subseteq \mathcal{S}$). Then,

$$\begin{aligned}\text{IG}(\mathcal{T}; X_m) &:= \mathbb{H}[x^*] - \mathbb{H}[x^*|X_m] \\ &= \frac{1}{2}\log v_* + \frac{1}{2}\log(2\pi e) - \left[\frac{1}{2}\log \sigma_m^2(x^*) + \frac{1}{2}\log(2\pi e)\right] \\ &= \frac{1}{2}\log v_* - \frac{1}{2}\log \sigma_m^2(x^*) \\ &\leq \frac{1}{2}\log v_* - \frac{1}{2}\log(k_* - M_*)\end{aligned} \tag{6}$$

where $\mathbb{H}[\cdot]$ is differential entropy and the last inequality will be derived subsequently. Intuitively, if observing X_m reduces $\sigma_m^2(x^*)$ (i.e., having a low conditional entropy $\mathbb{H}[x^*|X_m]$), then we are more certain about x^*, which leads to a larger IG.

Now we analyse the expressions in Eq. (6) and their relationships with X_m and $\mathcal{S}, \mathcal{T}$. First, v_* only depends on the prior on x^* (i.e., the specific GP instance including the choice for kernel and its hyperparameters and the noise variance λ), and is independent on X_m so we can treat it as a constant for analysing the effect of X_m on IG. We can obtain the following lower bound on $\sigma_m^2(x^*)$ by using the assumed condition on the relationship between x^* and X_m (or $\mathcal{S}$): $\inf_{x\in\mathcal{S}} \|x - x^*\|_2 \geq \delta$ which implies $\min_{x\in X_m} \|x - x^*\|_2 \geq \delta$. It means agent i is constrained in the data acquisition capabilities and unable to acquire data arbitrarily close to the target x^*. Recall that a low posterior variance $\sigma_m^2(x^*)$ means agent i is certain about the target x^* and learning is effective, so a lower bound on $\sigma_m^2(x^*)$ translates to an upper bound on IG/learning effectiveness.

Precisely, define $k_m(x^*) := \{\mathcal{K}(x_l, x^*)\}_{x_l \in X_m}$,

$$\begin{aligned}\sigma_m^2(x^*) &\triangleq \mathcal{K}(x^*, x^*) - k_m(x^*)^\top \widetilde{K}_{X_m}^{-1} k_m(x^*) \\ &= k_* - k_m(x^*)^\top \widetilde{K}_{X_m}^{-1} k_m(x^*) \\ &\geq k_* - \lambda_{\max}[\widetilde{K}_{X_m}^{-1}] \times k_m(x^*)^\top k_m(x^*) \\ &= k_* - \lambda_{\min}[\widetilde{K}_{X_m}] \times k_m(x^*)^\top k_m(x^*) \\ &\geq k_* - \underbrace{\left[\text{tr}(\widetilde{K}_{X_m}) - (m-1)\left(\frac{m\det(\widetilde{K}_{X_m})}{\text{tr}(\widetilde{K}_{X_m})}\right)^{1/(m-1)}\right] \times (m\varepsilon_{\mathcal{K}}^2)}_{M_*}\end{aligned}$$

where $\lambda_{\max}[\cdot]$ finds the maximum eigenvalue of a matrix (similarly for $\lambda_{\min}$). The two terms in M_* are derived respectively next.

The third line (first inequality) is due to Lemma 3. The fourth line is due to the relationship between eigenvalues of a squared matrix and those of its inverse (i.e., $\lambda_i \neq 0$ is a valid eigenvalue for a matrix A iff $1/\lambda_i$ is a valid eigenvalue for A^{-1}) and all eigenvalues for the positive definite $\widetilde{K}_{X_m}$ are positive. Next, we derive and interpret the two terms in M_* in the last line, respectively.

The first term is an upper bound on the smallest eigenvalue of a positive (semi-)definite hermitian matrix $\widetilde{K}_{X_m}$ (Kumar, 1984). We note the following implication. If the variance of the observation noise λ is larger, this upper bound on $\lambda_{\min}$) is looser (larger) because the trace increases. Consequently, the lower bound for $\sigma_m^2(x^*)$ is looser (smaller), and the upper bound on IG is looser (larger). It can lead to an overly optimistic estimate of the learning effectiveness. In other words, if the observation noise is large, the IG can be over-estimated.

The second term is by exploiting the (decreasing) monotonicity of $\mathcal{K}(x^*,\cdot)$ in $\|x^*-\cdot\|_2$. Explicitly using the observational constraint, define $\varepsilon_{\mathcal{K}} = \varepsilon_{\mathcal{K}}(x^*) \coloneqq \sup_{x\in\mathcal{S}} \mathcal{K}(x^*,x)$. Then, $k_m(x^*)^\top k_m(x^*) = \sum_{x\in X_m} \mathcal{K}(x^*,x)^2 \leq t\varepsilon_{\mathcal{K}}^2$. Given a specific kernel such as squared exponential, we can derive the monotonic relationship between $\mathcal{K}(x^*,\cdot)$ and $\|x^*-\cdot\|_2$ to give a specific expression of $\varepsilon_{\mathcal{K}}$ using δ (and other kernel-specific hyperparameters). This analysis sheds light on a fundamental limit to learning effectiveness if all acquired data (or the entire support) are outside of the δ-ball centred at x^*. In other words, if the agents are constrained by their resources (so that they cannot observe arbitrarily close to x^*), there is an upper bound on the effectiveness of learning (i.e., IG). □

Proof of Proposition 2. Substitute Eq. (8) into Eq. (1),

$$\mathrm{IG}(\mathcal{T};X_m) = \frac{1}{2}\log\det\Sigma_{\mathcal{T}} - \frac{1}{2}\log(\det(\Sigma_{\mathcal{T}})\times\widetilde{d}_*) = -\frac{1}{2}\log\widetilde{d}_* \ .$$

This greatly simplifies the derivation to instead focus on a single term which is a (log) determinant.

Focusing on the matrix (whose determinant is denoted as $\widetilde{d}_*$) $\mathbf{I}_m - K_{X_m\mathcal{T}}K_{\mathcal{T}}^{-1}K_{\mathcal{T}X_m}\widetilde{K}_{X_m}^{-1}$, we can provide an upper bound on $\mathrm{IG}(\mathcal{T};X_m)$ by providing a lower bound on $\widetilde{d}_*$, which is achieved by bounding $\underbrace{K_{X_m\mathcal{T}}K_{\mathcal{T}}^{-1}K_{\mathcal{T}X_m}}_{A}\underbrace{\widetilde{K}_{X_m}^{-1}}_{B}$ to be close to the zero matrix $\mathbf{0}_m$ via $|AB[p,p]|\leq\xi_{\text{diag}}$ and $|AB[p,q]|\leq\xi_{\text{off}}, p\neq q$ for some small $\xi_{\text{diag}},\xi_{\text{off}}>0$.

Apply Lemma 4 for a sharp lower bound,

$$\widetilde{d}_* \geq (1-\xi_{\text{diag}}-(t-1)\xi_{\text{off}})(1-\xi_{\text{diag}}+\xi_{\text{off}})^{t-1} \ . \tag{7}$$

Now for $\xi_{\text{diag}},\xi_{\text{off}}$. Denote $\psi_p^\top \coloneqq [\mathcal{K}(x_p,x_1^*),\ldots,\mathcal{K}(x_p,x_{m'}^*)]$ for convenience. Then,

$$\begin{aligned}
AB[p,q] &\triangleq \langle \mathrm{row}_A(p), \mathrm{col}_B(q)\rangle \\
&= \langle [\psi_p^\top K_{\mathcal{T}}^{-1}\psi_1,\ldots,\psi_p^\top K_{\mathcal{T}}^{-1}\psi_m], \mathrm{col}_B(q)\rangle \\
&= \langle [\psi_p^\top K_{\mathcal{T}}^{-1}\psi_1,\ldots,\psi_p^\top K_{\mathcal{T}}^{-1}\psi_m], [\widetilde{K}_{X_m}^{-1}[q,p],\ldots,\widetilde{K}_{X_m}^{-1}[q,m]\rangle \\
&= \sum_{k=1}^m \psi_p^\top K_{\mathcal{T}}^{-1}\psi_k \times \widetilde{K}_{X_m}^{-1}[q,k] \\
&\leq \sum_{k=1}^m \lambda_{\max}[K_{\mathcal{T}}^{-1}]\psi_p^\top\psi_k \times \widetilde{K}_{X_m}^{-1}[q,k] && \text{(Lemma 3)} \\
&= \sum_{k=1}^m \lambda_{\min}[K_{\mathcal{T}}]\psi_p^\top\psi_k \times \widetilde{K}_{X_m}^{-1}[q,k] && \text{(Inverse eigenvalues for a positive definite matrix)} \\
&= \lambda_{\min}[K_{\mathcal{T}}]\sum_{k=1}^m \psi_p^\top\psi_k \times \widetilde{K}_{X_m}^{-1}[q,k] \\
&= \lambda_{\min}[K_{\mathcal{T}}]\sum_{k=1}^m (\sum_{l=1}^{m'} \mathcal{K}(x_p,x_l^*)\times\mathcal{K}(x_k,x_l^*)) \times \widetilde{K}_{X_m}^{-1}[q,k] \\
&\leq \lambda_{\min}[K_{\mathcal{T}}]\sum_{k=1}^m (\sum_{l=1}^{m'} \varepsilon_1\times\varepsilon_1) \times \widetilde{K}_{X_m}^{-1}[q,k] && \text{(Using the assumption } \sup_{x,x^*}\mathcal{K}(x,x^*)\leq\varepsilon_1)
\end{aligned}$$

$$
\begin{aligned}
&= \lambda_{\min}[K_{\mathcal{T}}] \sum_{k=1}^{m} (m\varepsilon_1^2) \times \widetilde{K}_{X_m}^{-1}[q,k] \\
&= \lambda_{\min}[K_{\mathcal{T}}] m' \varepsilon_1^2 \sum_{k=1}^{m} \widetilde{K}_{X_m}^{-1}[q,k] \\
&\leq \lambda_{\min}[K_{\mathcal{T}}] m' \varepsilon_1^2 \sum_{k=1}^{m} \widetilde{K}_{X_m}^{-1}[q,k] && \text{(Using the assumption } \max_{j,k} \widetilde{K}_{X_m}^{-1}[q,k] \leq \varepsilon_2\text{)} \\
&= \lambda_{\min}[K_{\mathcal{T}}] m' \varepsilon_1^2 \sum_{k=1}^{m} \varepsilon_2 \\
&= \lambda_{\min}[K_{\mathcal{T}}] m m' \varepsilon_1^2 \varepsilon_2 \ .
\end{aligned}
$$

For generality, use a uniform bound $\max(\xi_{\text{diag}}, \xi_{\text{off}}) \leq \lambda_{\min}[K_{\mathcal{T}}] m m' \varepsilon_1^2 \varepsilon_2$. Substituting this into Eq. (7) completes the proof.

For a non-vacuous bound, as IG is inherently upper bounded by the prior uncertainty (entropy),

$$
\text{IG}(\mathcal{T}; X_m) \leq 1/2 \min[\log |\Sigma_{\mathcal{T}}| + m \log(2\pi e), (1-m) \log((1 - \xi_{\text{diag}} - (t-1)\xi_{\text{off}})(1 - \xi_{\text{diag}} + \xi_{\text{off}}))] \ ,
$$

but this is not necessary for our other analysis.

The existence of ε_1 is due to the monotonicity of kernel $\mathcal{K}$ in $\|x - x^*\|_2$ (i.e., as δ increases ε_1 monotonically decreases) and it can be explicitly derived w.r.t. a specific choice of kernel. For example, for the radial basis kernel $\mathcal{K}_{\text{rbf}}(\|x - x'\|_2) := \exp(-\|x - x'\|_2^2/(2\sigma^2))$, $\varepsilon_1(\delta, \mathcal{K}_{\text{rbf}}) = \exp(-\delta^2/(2\sigma^2))$. The value of ε_2 represents the largest (regularised) partial correlation between any two observations x_q, x_w, conditioned on all the other observations $x_p, p \neq q, p \neq w$. It is small because typical active learning methods rarely acquire extremely close data (in Euclidean distance), and as a result, any two different data have small partial correlations (conditioned on other data).

□

A.4 Derivation of Eq. (3)

Proof. Apply Lemma 5 to $\det(\Sigma_{\mathcal{T}|X_m})$ in IG in Eq. (1),

$$
\widetilde{d}_* := \det(\Sigma_{\mathcal{T}|X_m}) / \det(K_{\mathcal{T}}) = \det(\mathbf{I}_t - K_{X_m\mathcal{T}} K_{\mathcal{T}}^{-1} K_{\mathcal{T}X_m} \widetilde{K}_{X_m}^{-1}) \ , \tag{8}
$$

combined with $K_{\mathcal{T}} = \Sigma_{\mathcal{T}}$ (by definition),

$$
\text{IG}(\mathcal{T}; X_m) = 1/2[\log\det \Sigma_{\mathcal{T}} - \log(\det(\Sigma_{\mathcal{T}}) \times \widetilde{d}_*)] = -1/2 \log \widetilde{d}_* \ .
$$

□

Observations based on $\tilde{d}_*$ and IG upper bound for general $\mathcal{T}$. Based on Eq. (8) and Eq. (3), we make two observations: (i) in Eq. (1), $\text{IG}(\mathcal{T}; X_m) \geq 0$, and since $\Sigma_{\mathcal{T}} = K_{\mathcal{T}}$, it implies $\widetilde{d}_* \in [0, 1]$ where $\widetilde{d}_* = 0$ corresponds to *exact* learning (i.e., IG is maximum) and $\widetilde{d}_* = 1$ corresponds to no learning at all (i.e., $\text{IG}(\mathcal{T}; X_m) = 0$). To see its implication, suppose perfect learning capabilities with no observation noise, i.e., $\lambda = 0$ (so $\widetilde{K}_{X_m} = K_{X_m}$) and perfect acquisition of data $\mathcal{T} = X_m$, then we have

$$
\mathbf{I}_m - K_{X_m\mathcal{T}} K_{\mathcal{T}}^{-1} K_{\mathcal{T}X_m} \widetilde{K}_{X_m}^{-1} = \mathbf{I}_m - K_{\mathcal{T}} K_{\mathcal{T}}^{-1} K_{\mathcal{T}} K_{\mathcal{T}}^{-1} = \mathbf{0}_m
$$

which gives $\widetilde{d}_* = 0$ (exact learning).[5]

[5] In this scenario, if we substitute $\mathbf{0}_m$ back into Eq. (1), it becomes undefined since the determinant is 0 and $\log 0$ is undefined. This is due to the fact that any covariance matrix (e.g., $\Sigma_{\mathcal{T}|X_m}$) needs to be positive (semi-)definite and, in particular invertible. This requirement is violated in exact learning since observing X_m without any noise already completely tells us about $\mathcal{T} = X_m$. We describe this special and extreme case solely for illustrative purposes (to describe exact learning). Our following analysis will stipulate to $\Sigma_{\mathcal{T}|X_m}$ being invertible and so $\log\det \Sigma_{\mathcal{T}|X_m}$ in Eq. (1) is always well-defined.

(ii) The term $\det(K_{\mathcal{T}})$ is independent of X_m, so we focus on $\widetilde{d}_*$ for analyzing learning effectiveness w.r.t. X_m (and $\mathcal{S}, \mathcal{T}$). Its implication is: Suppose there is noise i.e., $\lambda > 0$ but with perfect acquisition of data $\mathcal{T} = X_m$, then

$$\mathbf{I}_m - K_{X_m\mathcal{T}}K_{\mathcal{T}}^{-1}K_{\mathcal{T}X_m}\widetilde{K}_{X_m}^{-1} = \mathbf{I}_m - K_{X_M}\widetilde{K}_{X_m}^{-1}$$

and its determinant $\widetilde{d}_*$ is small (i.e., learning is effective) if the noise variance λ is small.

These two illuminating observations are under ideal conditions: perfect/noise-free observation and/or perfect data acquisition capability $\mathcal{T} = X_m$. In practice, perfect (noise-free) observation is improbable (especially for physical, (bio)chemical and material science experiments), which is often modelled via an additive (zero-mean) Gaussian noise with variance λ Rasmussen and Williams (2006) as mentioned in §2. More importantly, the main motivation for CAL is that each agent does *not* have perfect acquisition (i.e., $\inf_{x\in\mathcal{S},x^*\in\mathcal{T}} \|x - x^*\|_2 \geq \delta$).

Comparative discussion on explicitly specifying targets vs. learning about f in general. In our setting, AL is used to collect data to learn about (the functional mapping) on some explicitly specified target $\mathcal{T}$. In particular, the functional mapping is w.r.t. the ground truth function f, the observations of which contain an additive zero-mean Gaussian noise with variance λ, as in Sec. 2.

One primary reason for this setting is so that the IG w.r.t. an explicitly specified target $\mathcal{T}$ has a closed-form expression in GP. This enables the subsequent analysis and derivations of individual rationality and fairness (e.g., Proposition 3 and Proposition 4).

We note that while the practical use cases investigated in our experiments, such as the material design and drug discovery, might also be addressed with an alternative setting: each agent is interested in learning the functional mapping, i.e., f itself instead of $\mathcal{T}$, it may not enable analysis with closed-form expressions.[6] In addition, to learn about f itself, it means both the input space and output domain of f may need to be fixed (e.g., the output domain for regression is $\mathbb{R}$). However, our setting circumvents this need to explicitly fix the output domain of f as the IG calculation (and subsequently Eq. (2)) only requires the input space $\mathcal{X}$ to be fixed (and a suitable kernel is available). This implies our method is more flexible. For instance, in our experiments, while the output space of f is $\mathbb{R}$ for DD and MD, the output space of f for DED is $\mathbb{R}^2$ as it is w.r.t. a 2-dimensional differential equation (where we have used the vector-valued operator kernel (Heinonen et al., 2018)). In contrast, it may be challenging to apply a method which assumes the output space of the f to be $\mathbb{R}$ to another problem where the output space is different such as $\mathbb{R}^2$.

Lastly, our setting of explicitly specified targets can also be used to address aforementioned use cases via a perspective of a discrete approximation of the continuous input space.[7] In particular, learning about the function f, it is equivalent to learning the functional values $f(x)$ over the entire input space $\forall x \in \mathcal{X}$. However, for a continuous input space $\mathcal{X}$, it is not tractable in practice. Therefore, a discretisation of $\mathcal{X}$ is often adopted, and $\mathcal{T}$ can be viewed as such discretisation, except it can additionally encode an agent's preference about which part of the input space. In other words, if the agent wants to learn about the functional values of the entire input space in general, it can specify its $\mathcal{T}$ to contain uniformly "spread out" input locations in $\mathcal{X}$. An example of this is investigated via the DED dataset, where each agent is interested in learning about the *entire* underlying differential equation through GP regression as the modelling framework. In this case, their respective (similar but different) targets are over the entire input space (i.e., the time dimension). We highlight that this is to show our setting can address the use-cases where the agents are interested in the functional mapping in general by specifying their targets accordingly, and our work does *not* discuss *how* such targets are specified, but treats them as given.

In summary, our setting of explicitly specifying the targets has the following advantages over learning about the function f in general: **(1)** it enables closed-form expressions for IGs and tractable analysis; **(2)** it is more flexible in *not* having to fix the output space of f; and **(3)** it can be seen as a discrete approximation of learning about the function f in general.

A.5 Proof of Lemma 1

Proof. Given a set X of m data, recall the expression for $\mathrm{IG}(\mathcal{T}; X)$,

$$\mathrm{IG}(\mathcal{T}_i; X) = -1/2 \log\det(\mathbf{I}_m - K_{X\mathcal{T}}K_{\mathcal{T}}^{-1}K_{\mathcal{T}X}\widetilde{K}_X^{-1}) \,,$$

and define auxiliary matrices

$$D \coloneqq K_{X\mathcal{T}} - K_{\mathcal{T}}, \quad E \coloneqq K_{\mathcal{T}} - \widetilde{K} + \lambda\mathbf{I} \,.$$

[6] Daxberger and Low (2017) provides a formulation (i.e., precise expression of IG w.r.t. the function f itself) which is amenable to the alternative: directly learning f instead of some explicitly specified targets. It would be an interesting future direction.

[7] Note that if the input space itself is discrete, then our setting follows naturally where $\mathcal{T}$ is a subset of the discrete input space.

Substitute the auxiliary matrices into $\mathbf{I}_m - K_{X\mathcal{T}}K_{\mathcal{T}}^{-1}K_{\mathcal{T}X}\widetilde{K}_X^{-1}$,

$$\begin{aligned}\mathbf{I}_m - K_{X\mathcal{T}}K_{\mathcal{T}}^{-1}K_{\mathcal{T}X}\widetilde{K}_X^{-1} &= (K_{\mathcal{T}}+D)K_{\mathcal{T}}^{-1}(K_{\mathcal{T}}+D)\widetilde{K}_X^{-1}\\ &= (K_{\mathcal{T}}+2D+D^2)\\ &= \mathbf{I}+(2D+D^2+E)\widetilde{K}_X^{-1}\,.\end{aligned}$$

Consequently,

$$\begin{aligned}\mathrm{IG}(\mathcal{T};X) &= -\frac{1}{2}\log\det(\mathbf{I}-\mathbf{I}+(2D+D^2+E)\widetilde{K}_X^{-1}\\ &= -\frac{1}{2}\log\det((2D+D^2+E)(\widetilde{K}_X^{-1}))\\ &= -\frac{1}{2}\log(\det((2D+D^2+E)\det(\widetilde{K}_X^{-1}))\,. \end{aligned} \tag{9}$$

Bound each diagonal entry of the matrix $(2D+D^2+E)$ to bound its trace and determinant:

$$\begin{aligned}&(2D+D^2+E)[pp] \triangleq 2D[pp]+D^2[pp]+E[pp]\\ &= 2\underbrace{(\mathcal{K}(x_p,x_p^*)-\mathcal{K}(x_p^*,x_p^*))}_{D[pp]}+\underbrace{\sum_{q=1}^{m'}(\mathcal{K}(x_p,x_q^*)-\mathcal{K}(x_p^*,x_q^*))^2}_{D^2[pp]}+\underbrace{\mathcal{K}(x_p,x_p^*)-\mathcal{K}(x_p,x_p)-\lambda}_{E[pp]}\\ &= \mathcal{K}(x_p,x_p^*)-\mathcal{K}(x_p^*,x_p^*)+\mathcal{K}(x_p,x_p^*)-\mathcal{K}(x_p,x_p)-\lambda+\sum_{q=1}^{m'}(\mathcal{K}(x_p,x_q^*)-\mathcal{K}(x_p^*,x_q^*))^2\\ &\le \bar{\varepsilon}+\bar{\varepsilon}+m\bar{\varepsilon}^2-\lambda\,.\end{aligned}$$

Use the result $\det(A)\le(\frac{\mathrm{tr}(A)}{m})^m, A\in\mathbb{R}^{m\times m}$,

$$\det(2D+D^2+E)\le(2\bar{\varepsilon}+m\bar{\varepsilon}-\lambda)^m\,.$$

Substituting this back into Eq. (9) completes the proof.

□

A.6 Proof of Proposition 3

Definition 2 (Covering Set for $\mathcal{T}_i$). For target $\mathcal{T}_i=\{x_1^*,\dots,x_{m'}^*\}$, a covering set $X_{\text{cover},i}=\{x_1,\dots,x_m\}$ is s.t.

$$\bar{\varepsilon}\coloneqq\max_{x_l\in X_{\text{cover},i},x_{l'}^*\in\mathcal{T}_i}\sup_{x\in\mathcal{X}}|\mathcal{K}(x_l,x)-\mathcal{K}(x_{l'}^*,x)|\le\left(\lambda+\sqrt[m']{\frac{\det\widetilde{K}_X}{\varepsilon_{\text{diag}}^{(1-m)/(1-1/e)}}}+\frac{1}{m'^2}\right)^{1/2}-\frac{1}{m'}$$

where $\varepsilon_{\text{diag}}\coloneqq(1-\xi_{\text{diag}}-(m-1)\xi_{\text{off}})(1-\xi_{\text{diag}}+\xi_{\text{off}})$ and $\xi_{\text{diag}},\xi_{\text{off}}$ are defined in Proposition 2.

By combining Proposition 2 and Lemma 1, Definition 2 describes the condition on the set $X_{\text{cover},i}$ of data (i.e., covering set) that can provide a higher IG than $\overline{\text{IG}}_{i,\text{indiv}}$, which leads to Proposition 3.

Proof of Proposition 3. This is by directly substituting the definition of a covering set and verifying the following inequalities.

$$\overline{\text{IG}}_{i,\text{indiv}}\le(1-1/e)\text{IG}_{\text{cover}}\le(1-1/e)\text{IG}(\mathcal{T}_i;D_{N,m}^*)\le\text{IG}(\mathcal{T}_i;D_{N,m}) \tag{10}$$

where $\text{IG}_{\text{cover}}\coloneqq\text{IG}(\mathcal{T}_i;X)$ and $\overline{\text{IG}}_{i,\text{indiv}}$ denotes the IG upper bound for without collaboration from Proposition 2.

This is a constructive proof: substitute the values of $\bar{\varepsilon}$ to verify the first inequality in Eq. (10). Lemma 1 is used.

The second inequality is due to the definition of $D_{N,m}^*$ which is optimal so that $\text{IG}_{\text{cover}}\le\text{IG}(\mathcal{T}_i;D_{N,m}^*)$ and the last inequality is due to Lemma 6. □

A.7 Derivation of Claim 1

Derivation of Claim 1. Note that $\text{IG}_{i,0} = 0, \forall i \in N$. Given some small $\delta_{\text{IG}} > 0$ as a specified requirement on the equitability of the IGs.

Substitute $\log(\rho_{i,k}) = \Delta_{i,k}$,

$$\max\left(\prod_{k=1}^{m}\frac{\rho_{i,k}}{\rho_{j,k}}, \prod_{k=1}^{m}\frac{\rho_{j,k}}{\rho_{i,k}}\right) \leq \exp(\delta_{\text{IG}}) \implies |\text{IG}_{i,m} - \text{IG}_{j,m}| \leq \delta_{\text{IG}} \,.$$

Relaxing the condition in Claim 1. As mentioned after Proposition 4, the condition in Claim 1

$$\text{for } 1 \leq k \leq m, \max\left(\frac{\rho_{i,k}}{\rho_{j,k}}, \frac{\rho_{j,k}}{\rho_{i,k}}\right) \leq \exp(\delta_{\text{IG}})/m \,,$$

can be relaxed to

$$\max\left(\prod_{k=1}^{m}\frac{\rho_{i,k}}{\rho_{j,k}}, \prod_{k=1}^{m}\frac{\rho_{j,k}}{\rho_{i,k}}\right) \leq \exp(\delta_{\text{IG}}) \,.$$

This relaxed condition also satisfies the implication that $|\text{IG}_{i,m} - \text{IG}_{j,m}| \leq \delta_{\text{IG}}$. □

A.8 Proof of Lemma 2

The high-level idea is by finding the property, specifically, w.r.t. ε-PDP (Definition 1) of the SWO defined as in Eq. (2) with a set of sharing coefficients $\{\beta_i\}_{i\in N}$.

$\zeta_{\bar{\beta}}$, the maximum over all pairwise differences (relative to some β_i), characterises how different these sharing coefficients are. $\varepsilon_{i,j}$, derived from the condition between $i, j \in N$ from the ε-PDP definition, leads to an implication in ε-PDP when it is *not* vacuously true (i.e., ensures the condition is true), so as to derive the precise conditions (i.e., ε_β) for Eq. (2) to satisfy ε-PDP.

Proof. For Eq. (2) to satisfy ε-PDP, it requires $\forall \varepsilon' > \varepsilon, (\exists i, j \in N, \rho_i - \varepsilon' = \varrho_i \geq \varrho_j = \rho_j + \varepsilon') \wedge (\forall l \in N \setminus \{i, j\}, \rho_l = \varrho_l) \implies \boldsymbol{\rho} \preceq_{\text{Equ. (2)}} \boldsymbol{\varrho}$. Note that $\boldsymbol{\rho} \preceq_{\text{Equ. (2)}} \boldsymbol{\varrho}$ is equivalent to $\prod_i \rho_i^{1/\beta_i} \leq \prod_i \varrho_i^{1/\beta_i}$, which will be shown in the proof.

As ε does not affect $(\forall l \in N \setminus \{i, j\},\ \rho_l = \varrho_l)$, the proof focuses on $(\exists i, j \in N, \rho_i - \varepsilon' = \varrho_i \geq \varrho_j = \rho_j + \varepsilon')$ and uses $\rho_l = \varrho_l, \forall l \in N \setminus \{i, j\}$ to simplify the inequality $\prod_i \rho_i^{1/\beta_i} \leq \prod_i \varrho_i^{1/\beta_i}$ to $\rho_i^{1/\beta_i}\rho_j^{1/\beta_j} \leq \varrho_i^{1/\beta_i}\varrho_j^{1/\beta_j}$.

Note that if ε satisfies the following implication

$$(\rho_i - \varepsilon = \varrho_i \geq \varrho_j = \rho_j + \varepsilon) \implies \rho_i^{1/\beta_i}\rho_j^{1/\beta_j} \leq \varrho_i^{1/\beta_i}\varrho_j^{1/\beta_j} \,,$$

then all $\varepsilon' > \varepsilon$ also satisfies it, restricted to $\varepsilon' \in (\varepsilon, |\rho_i - \rho_j|/2]$. The upper bound $|\rho_i - \rho_j|/2$ is necessary to preserve the correct order $\rho_i - \varepsilon' \geq \rho_j + \varepsilon'$ in PDP.

It remains to show ε satisfies this implication. The proof first simplifies the inequality, beginning with raising both sides to the power of β_j,

$$1^{\beta_j} = 1 \leq \frac{\varrho_i\varrho_j}{\rho_i\rho_j}(\rho_i\rho_j)^{\frac{\beta_j-\beta_i}{\beta_i}} \,.$$

Substitute $\varrho_i = \rho_i - \varepsilon$ and $\varrho_j = \rho_j + \varepsilon$,

$$1 \leq \frac{(\rho_i - \varepsilon)(\rho_j + \varepsilon)}{\rho_i\rho_j}(\rho_i\rho_j)^{\frac{\beta_j-\beta_i}{\beta_i}} \,.$$

Collect $\rho_i\rho_j$ to one side,

$$(\rho_i - \varepsilon)(\rho_j + \varepsilon) \geq (\rho_i\rho_j)^{1+\frac{\beta_i-\beta_j}{\beta_i}} \,.$$

Substitute the definition of $\varepsilon_{i,j}$,

$$(\rho_i - \varepsilon_{i,j})(\rho_j + \varepsilon_{i,j}) \geq (\rho_i \rho_j)^{1+\zeta_{\bar{\beta}}} \geq (\rho_i \rho_j)^{1+\frac{\beta_i - \beta_j}{\beta_i}} .$$

Substituting the definition of ε_β completes the proof. □

Note on the social welfare ordering ι. Typically ι is ordinal: ι gives an ordering between two $\boldsymbol{\rho}, \boldsymbol{\varrho}$. A different version called cardinal SWO orders any two $\boldsymbol{\rho}, \boldsymbol{\varrho}$ by calculating (and comparing) some values. For instance, NSW is a cardinal SWO by comparing $\prod_i \rho_i$ vs. $\prod_i \varrho_i$ (larger is preferred). A cardinal SWO is more "informative" in the sense that there is some numerical value of each $\boldsymbol{\rho}$ available (in addition to being used for ordering).

A.9 Proof of Proposition 4

Proof. For some $m \geq 1$,

$$\boldsymbol{\rho}_m = \rho_{1,1}, \ldots, \rho_{n,1}, \rho_{1,2}, \ldots, \rho_{n,2}, \ldots, \rho_{1,m}, \ldots, \rho_{n,m} ,$$

and

$$\boldsymbol{\varrho}_m = \varrho_{1,1}, \ldots, \varrho_{n,1}, \varrho_{1,2}, \ldots, \varrho_{n,2}, \ldots, \varrho_{1,m}, \ldots, \varrho_{n,m} .$$

From Lemma 2, $1 \leq k \leq m$,

$$\begin{aligned} \left(\rho_{i,k} - \zeta_{\bar{\beta},l} = \varrho_{i,k} \geq \varrho_{j,k} = \rho_{j,k} + \zeta_{\bar{\beta},l}\right) \bigwedge (\forall l \in N \setminus \{i,j\}, \rho_{l,k} = \varrho_{l,k}) \\ \implies \rho_{i,k} \times \rho_{j,k} \leq \varrho_{i,k} \times \varrho_{j,k} . \end{aligned}$$

Substitute the definition of $\varepsilon_{\text{global}}$ to verify $\varepsilon_{\text{global}}$-PDP,

$$\begin{aligned} (\rho_{i,k} - \varepsilon_{\text{global}} = \varrho_{i,k} \geq \varrho_{j,k} = \rho_{j,k} + \zeta_{\text{global}}) \bigwedge (\forall l \in N \setminus \{i,j\}, \rho_{l,k} = \varrho_{l,k}) \\ \implies \rho_{i,k} \times \rho_{j,k} \leq \varrho_{i,k} \times \varrho_{j,k} \\ \implies \prod_{k=1}^{m} \rho_{i,k} \times \rho_{j,k} \leq \prod_{k=1}^{m} \varrho_{i,k} \times \varrho_{j,k} . \end{aligned}$$

□

A.10 An Overview of Assumptions and Limitations of Presented Theoretical Results

We discuss at a high level some important assumptions and limitations of our presented theoretical results and refer the readers to the precise and specific conditions of the respective result (and their proofs/derivations).

Note on the conditional independence assumption used by Lemma 6. As an example of how the conditional independence assumption can be satisfied: if for each agent i , $\mathcal{S}_i \subseteq \mathcal{T}_i$, then the conditional independence is satisfied. While this may seem counter-intuitive to the motivation that each agent has limited ability to observe its target directly, one way to consider is this: Although $\mathcal{S}_i \subseteq \mathcal{T}_i$, agent i can still benefit from collaboration as $\mathcal{T}_i$ may contain additional input locations $\mathcal{U}_i$ such that $\mathcal{T}_i = \mathcal{S}_i \cup \mathcal{U}_i$. As agent i may have a limited budget and may not have access to collect observations near $\mathcal{U}_i$, other agents can provide additional observations, particularly observations that are highly informative to $\mathcal{U}_i$ (which help agent i) while agent i can, in turn, provide such observations to other agents, corresponding to the *quid pro quo* use-case highlighted to Sec. 1.

Our results (Proposition 1, Proposition 2) defined δ to be the minimum distance between any point x in the support $\mathcal{S}_i$ and any point x^* in the target $\mathcal{T}_i$ of the agent i in order to clearly illustrate the observational constraint, so it may seem counter-intuitive to consider the scenario of $\mathcal{S}_i \subset \mathcal{T}_i$. We clarify here that the formalisation of observational constraint using δ is primarily to clearly illustrate the observational constraint and to aid the interpretation of the results subsequently. However, the specific value of δ does *not* have a precise and quantitative effect on the results (Proposition 1, Proposition 2). To elaborate, in Proposition 1, the upper bound of IG depends on $\varepsilon_{\mathcal{K}}$, which is a monotonically decreasing relationship with δ, but the precise relationship between $\varepsilon_{\mathcal{K}}$ and δ is not specified since the choice of the kernel is not fixed. Similarly, in

Proposition 2, ε_1 is monotonically decreasing in δ, but the precise relationship between them is not specified without having fixed the kernel. In this regard, the specific value of δ does not have a critical quantitative effect in the results Proposition 1 and Proposition 2. Then, the scenario of $\mathcal{S}_i \subset \mathcal{T}_i$ can correspond to a special case where $\delta = 0$ in both results. We highlight that this does *not* change the derivation or the result, but is a specific case to illustrate how conditional independence and thus the submodularity of IG in Eq. (1) can be satisfied, which is used later in Proposition 3 to derive individual rationality (IR).

As an *informal* intuition on how $\mathcal{S}_i \subset \mathcal{T}_i$ can lead to conditional independence: For any two disjoint proper subsets X_A, X_B of $\mathcal{S}_i$: $X_A, X_B \subset \mathcal{S}_i, X_A \cap X_B \emptyset$, we can consider the following conditional distributions $P(X_A|\mathcal{T}_i, X_B)$ and $P(X_A|\mathcal{T}_i)$. In particular, if they are the same, then by definition, we have conditional independence. To see why $P(X_A|\mathcal{T}_i, X_B) = P(X_A|\mathcal{T}_i)$, recall that $X_B \subset \mathcal{S}_i \subset \mathcal{T}_i$. Hence, $P(X_A|\mathcal{T}_i, X_B) = P(X_A|\mathcal{T}_i)$ as X_B is already contained in $\mathcal{T}_i$.

In our experiments, *without* having enforced the condition of $\mathcal{S}_i \subset \mathcal{T}_i$ (which means that submodularity of Eq. (1) and Eq. (2) are not automatically guaranteed), we still observe that IR is satisfied (e.g., Tables 3 and 7): the IG from collaboration for each agent is markedly higher than that without collaboration. It is an interesting future work to identify a possibly less restrictive condition for the submodularity of Eq. (1) than conditional independence.

Assumption on the maximum difference in kernel values in Lemma 1. Lemma 1 presents a contrasting view to Proposition 1 and Proposition 2, and characterises a sufficient condition on the lower bound of the learning/IG. The assumption on the maximum (of supremum) difference in the kernel values between any collected point (in X) and any point in target (in $\mathcal{T}_i$) intuitively describes "how good" this collected dataset X is for learning about the target $\mathcal{T}_i$. Specifically, the lower this maximum, the better X is.

Proposition 3 takes advantage of both Proposition 2 and Lemma 1. Importantly, another assumption is that the optimisation of Eq. (2) is "near-optimal" as in Lemma 6, which requires Eq. (1) to be submodular (discussed above).

Assumption on the equitability of the sharing coefficients in Lemma 2. Lemma 2 characterises the interaction among the sharing coefficients β_i's so that Eq. (2) satisfies the generalised PDP. Generally, if the β_i's are more equitable (i.e., their values are closer to each other), the better the generalised PDP. The definition of $\varepsilon_{i,j}$ in Lemma 2 follows from the condition in the definition in (generalised) PDP. Subsequently, Proposition 4 extends the result of Lemma 2 (which is w.r.t. one iteration k) to all iterations. The key assumption is that the global result (i.e., over all iterations) can be decomposed to the local results (i.e., each iteration), as in Eq. (5).

B ADDITIONAL EMPIRICAL SETTING DETAILS AND EMPIRICAL RESULTS

B.1 Additional Dataset Settings and Implementation

Synthetic-2D Identical setting. In addition to the mismatch setting in Sec. 4 to intuitively demonstrate why the *quid-pro-quo* benefit can be achieved, the identical setting in Fig. 4 can help demonstrate why the cost-saving benefit can be achieved. In particular, left 3 plots of Fig. 4 show that each agent is able to observe reasonably close to their respective targets (which are similar in this case). However, each has a very limited budget $m = 5$ to "cover"/learn effectively about their targets. Right 3 plots of Fig. 4 show that if the agents pool their resources (i.e., 5 budget from each) and coordinate in collecting the data, they are able to cover their targets more effectively.

Implementation details for real-world data. The (unnormalised) sharing coefficients corresponding to greedy_1-4 used in the experiments for the real-world data are as follows, for MD: [19,1,10], [18,2,10], [15,5,10] and [1,1,1]; for DD: [19,1,10], [18,2,10], [16,5,10] and [1,1,1]; for DED: [10,5,1], [10, 2, 1], [10, 1, 1] and [1,1,1]. The input spaces, kernel choices and corresponding references for these real-world data are in Table 6. The specific data preprocessing steps and kernel hyperparameters follow these implementations. for MD: https://github.com/rajak7/Bayesian_Optimization_Material_design; for DD: https://github.com/kexinhuang12345/DeepPurpose and for DED, https://github.com/cagatayyildiz/npde/. Note that DED is more complex because the output space of the true function f is $\mathbb{R}^2$, and we adopt the operator kernel (Heinonen et al., 2018) in our method.

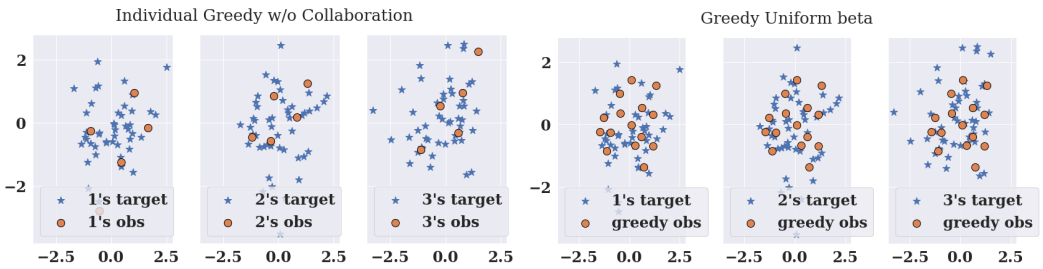

Figure 4: Identical: all agents have the same domain for their $\mathcal{T}_i, \mathcal{S}_i$. The point is that each agent can learn about their $\mathcal{T}_i$ reasonably well with *sufficient* budget. In this case, the budget is limited to 5 for each agent so the agents can benefit by pooling resources, i.e., their budgets.

Datasets	Input space $\mathcal{X}$	Kernel	Reference
MD	12-dimensional real value	Matérn kernel	Bassman et al. (2018)
DD	String	Gap-weighted subsequences kernel	Pahikkala et al. (2015)
DED	2-dimensional real value matched to time	Vector-valued operator kernel	Heinonen et al. (2018)

Table 6: GP implementation details for real-world datasets.

B.2 An Approximation for Computational Tractability

Recall Eq. (2) optimises over a Cartesian product in each iteration k, which can be computationally costly if n and/or $\mathcal{S}_i$ is large, so we describe an approximation.

Solving an approximate problem exactly. This approach reduces the exponential complexity to a polynomial through the way of approximation so that the polynomial problem can be solved exactly. Recall the exponential complexity in $\operatorname{argmax}_{\vec{x}_{N,k} \in \mathcal{S}_N}$ is due to the Cartesian product $\mathcal{S}_N \coloneqq \bigtimes_{i \in N} \mathcal{S}_i$. For notational simplicity, the subscript k is omitted (when considering a specific k) and denote $J(\vec{x}_N) \coloneqq \sum_i \Delta_i / \beta_i$.

Specifically, the approximate problem is finding the maximiser $\vec{x}_{N,M}$ in a uniformly sampled subset $\mathcal{S}' \subseteq \mathcal{S}_N$ of size M to approximate the true maximiser $\vec{x}^*_N \coloneqq \operatorname{argmax}_{\vec{x}_N \in \mathcal{S}_N} J(\vec{x}_N)$. In general $\mathcal{S}_N$ can be discrete or continuous depending on each $\mathcal{S}_i$.[8] The approximation analysis is w.r.t. the continuous case because it can be extended to the discrete case. W.l.o.g., consider $[0,1]^d$ as the input space (essentially a dimension-wise normalised version of $\mathcal{S}_N$ with d dimensions). Note that d is the dimension of the Cartesian product $\mathcal{S}_N$.

Theoretical guarantee. We describe the theoretical guarantee of this approximation approach via the relationship between a probabilistic error and the size of the subset $\mathcal{S}'$ in Corollary 1.

Corollary 1. For some small $\varepsilon_J > 0$, to achieve $J(\vec{x}^*_N) - J(\vec{x}_{N,M}) \leq \varepsilon_J$ w.p. $\geq 1 - \delta_J$, it requires $M \geq \ln(1/\delta_J) / \ln[1/(1 - V \times \varepsilon_J)]$ where $V \coloneqq \pi^{d/2} / [\Gamma(d/2+1) 2^{d+1} L]$.

Therefore, for an $(\varepsilon_J, \delta_J)$-approximation, the subset size M depends on $\varepsilon_J, 1/\delta_j$ logarithmically (the dependence on ε_J is less straightforward). The computational complexity is linear in M (assuming the computational complexity of $J(\cdot)$ is fixed): $M \times \mathcal{O}(J(\cdot))$. In this result, as L increases, M increases. We note that the analysis can be made more specific by obtaining a specific value for L w.r.t. a specific choice of kernel. Moreover, the belief of (the location of) $\vec{x}^*_N$ can be leveraged to improve the uniform random sampling. These interesting explorations are deferred to future work.

Proof of Corollary 1. Substitute ε_J, δ_J in to Lemma 7.

For ε_J,

$$L 2r \leq \varepsilon_J \implies r \leq \varepsilon_J / (2L) \,.$$

[8] We do not consider the hybrid case: some $\mathcal{S}_i$ is discrete while others are continuous.

For $\delta_J \in (0, 1/2]$,

$$\begin{aligned}
1-\delta_J &\geq 1-[1-V(\vec{x}^*_N, r)/2^d]^M \\
\delta_J &\leq [1-V(\vec{x}^*_N, r)/2^d]^M \\
\ln(\delta_J) &\leq M \ln(1-V(\vec{x}^*_N, r)/2^d) \\
M &\geq \frac{\ln(\delta_J)}{\ln(1-V(\vec{x}^*_N, r)/2^d)} \\
&= \frac{-\ln(\delta_J)}{-\ln(1-V(\vec{x}^*_N, r)/2^d)} \\
&= \ln(1/\delta_J) \Big/ \ln\left[\frac{1}{1-\frac{\pi^{d/2} r}{\Gamma(d/2+1)2^d}}\right] \\
&\geq \ln(1/\delta_J) \Big/ \ln\left[\frac{1}{1-\frac{\pi^{d/2} \varepsilon_J}{\Gamma(d/2+1)2^d 2L}}\right] .
\end{aligned}$$

Note that the volume of the d-sphere is $V(\vec{x}^*_N, r) = (\pi^{d/2} r)/\Gamma(d/2+1)$. The last inequality uses $r \leq \varepsilon_J/(2L)$. □

Lemma 7 (Maximiser Approximation in an r-Ball). For a search space $\mathcal{S} = [0,1]^d$ and an L-Lipschitz continuous function J: $\forall \vec{x}_N, \vec{x}'_N \in \mathcal{S}, |J(\vec{x}_N) - J(\vec{x}'_N)| \leq L\|\vec{x}_N - \vec{x}'_N\|$. For a uniformly randomly sampled subset $\mathcal{S}' \subseteq \mathcal{S}$ s.t., $M = |\mathcal{S}'|$, and denote the approximate maximiser $\vec{x}_{N,M} \coloneqq \text{argmax}_{\vec{x}\in\mathcal{S}'} J(\vec{x})$. Then, $J(\vec{x}^*_N) - J(\vec{x}_{N,M}) \leq 2Lr$ w.p. $\geq 1-[1-(V(\vec{x}^*_N, r)/2^d)]^M$ where $V(x, r')$ is the volume of the d-sphere $\text{B}_{r'}(x)$ centered at x with radius r.

Proof. The probability of a uniformly randomly sampled $\vec{x} \sim \mathcal{S}$ being in the d-sphere $\text{B}_r(\vec{x}^*_N)$ with radius $r \leq 1$ centered at $\vec{x}^*_N$,

$$\begin{aligned}
\text{Pr}_{\vec{x}\sim\mathcal{S}}[\vec{x} \in \text{B}_r(\vec{x}^*_N)] &= \text{Vol}(\text{B}_r(\vec{x}^*_N) \cap [0,1]^d)/\text{Vol}([0,1]^d) \\
&\geq \text{Vol}(\text{B}_r(\vec{x}^*_N)) \times (1/2)^d .
\end{aligned}$$

The inequality is when $\vec{x}^*_N$ is lies on the corner of $[0,1]^d$ so the volume of the intersection is the smallest (imagine in the for $d=2$, $\vec{x}^*_N = (0,0)$ or $\vec{x}^*_N = (0,1)$). In particular, $\text{Vol}([0,1]^d) = 1$.

Hence,

$$\text{Pr}[\vec{x} \notin \text{B}_r(\vec{x}^*_N), \forall \vec{x} \in \mathcal{S}'] \leq [1-(\text{Vol}(\text{B}_r(\vec{x}^*_N) \times (1/2)^d)]^M .$$

Take the compliment,

$$\text{Pr}[\vec{x} \in \text{B}_r(\vec{x}^*_N), \exists \vec{x} \in \mathcal{S}'] \geq 1-[1-(\text{Vol}(\text{B}_r(\vec{x}^*_N) \times (1/2)^d)]^M .$$

Denote such observation with $\vec{x}'$, and apply the Lipschitz condition,

$$J(\vec{x}^*_N) - J(\vec{x}_{N,M}) \leq J(\vec{x}^*_N) - J(\vec{x}') \leq L\|\vec{x}^*_N - \vec{x}'\| \leq L(\|\vec{x}^*_N\| + \|\vec{x}'\|) \leq L2r .$$

□

B.3 Additional Empirical Results

Learning performance and fairness for Synthetic-1D and DED-SDE. Table 7 and Table 8 show that our method performs, in general, the best or most competitive for the Synthetic-1D dataset and the DED-SDE dataset. While for DED-SDE, our method does not perform the best in terms of fairness (i.e., std_{IG} or std_{MSE}), it (dynamic β) performs second best, which is consistent with the trade-off demonstrated in previous results that dynamic β trades-off some performance for better fairness.

Table 7: $\overline{\text{IG}}_m$ and std_m for **Synthetic-1D**.

Baselines	$\overline{\text{IG}}_m$	std_{IG}
ind	20.537 (4.777)	N.A.
rand	68.963 (0.413)	0.008 (0.006)
entropy	32.536 (1.088)	0.047 (0.033)
joint	71.093 (0.621)	0.012 (0.009)
greedy_1	81.847 (3.171)	0.055 (0.039)
greedy_2	81.884 (3.035)	0.052 (0.037)
greedy_3	82.423 (1.371)	0.024 (0.017)
greedy_4	**83.346 (0.402)**	**0.007 (0.005)**
dynamic β	81.876 (0.679)	0.012 (0.008)

Table 8: Learning performance (IG, MSE) and fairness evaluation (stds) for **DED-SDE**. Average (SE) over 5 independent trials is reported (in brackets).

Baselines	$\overline{\text{IG}}_m$	std_{IG}	$\overline{\text{MSE}}$	std_{MSE}
ind	-6.703 (0.092)	N.A.	5.179 (0.803)	N.A.
rand	5.303 (0.237)	**0.063 (0.045)**	4.766 (0.319)	1.433 (0.252)
entropy	4.433 (0.359)	0.115 (0.081)	4.772 (0.423)	**1.054 (0.076)**
joint	6.351 (0.681)	0.152 (0.107)	4.604 (0.283)	1.361 (0.299)
greedy_1	6.372 (0.706)	0.157 (0.111)	4.668 (0.369)	1.421 (0.204)
greedy_2	6.373 (0.845)	0.187 (0.133)	**4.330 (0.460)**	1.280 (0.240)
greedy_3	6.321 (0.953)	0.213 (0.151)	4.753 (0.254)	1.278 (0.210)
greedy_4	**6.445 (0.706)**	0.155 (0.109)	4.556 (0.418)	1.327 (0.296)
dynamic β	6.279 (0.492)	0.111 (0.078)	4.654 (0.266)	1.082 (0.207)

Cumulative performance evaluation. As noted after the proof of Lemma 6 in Appendix A, it is practically motivated to ensure an *iterative* data collection method has a high *cumulative* performance: the collected data $X_{N,k}$ lead to high IGs over the iterations k, instead of only after the final iteration. For evaluation, we plot and compare $(\sum_i \text{IG}_{i,k}/\beta_i)$ against the iterations k.

Fig. 5 verifies the high cumulative performance of our method in that the IGs increase the most quickly over the iterations. Moreover, Fig. 6 echos Fig. 5 and demonstrates that our method (in general) outperforms the batch AL method (i.e., joint in green) in terms of cumulative performance over the iterations. Fig. 7 also shows our method clearly outperforms rand and entropy while slightly outperforms joint. Lastly, the comparatively larger standard errors in Fig. 7 (than in Fig. 6) confirm DED is a more challenging dataset/task than MD and DD.

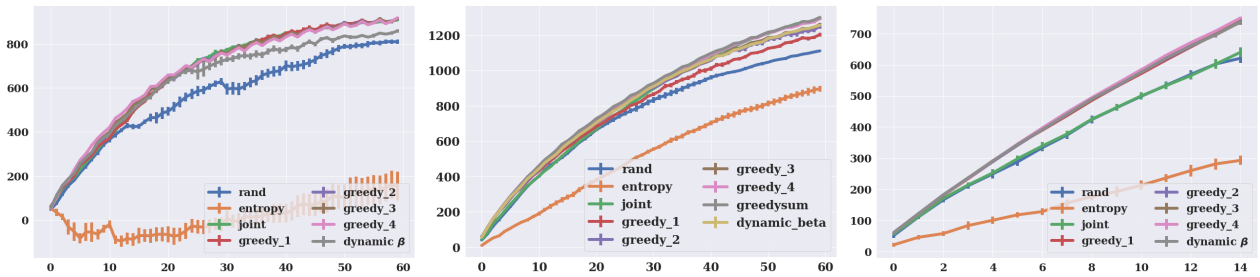

Figure 5: $\sum_i \text{IG}_{i,k}/\beta_i$ vs. iterations k on **Synthetic-1D** (left) **Synthetic-2D** mismatch (mid) & identical (right).

Vector field illustration of DED-VDP. Fig. 3 in Sec. 4 shows the comparison between the learning performance in DED-VDP of agent 1 without collaboration and with collaboration w.r.t. the 1st dimension of the DE. Fig. 8 plots the

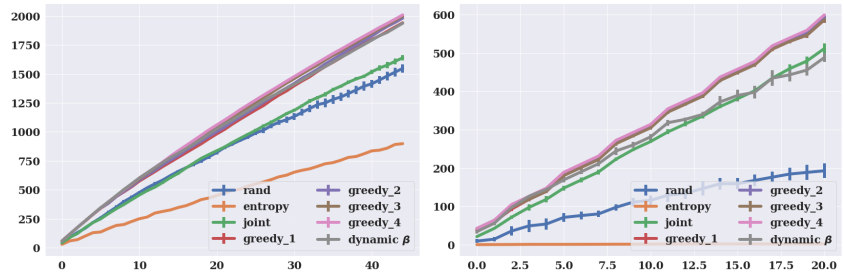

Figure 6: $\sum_i \mathrm{IG}_{i,k}/\beta_i$ vs. iterations k on MD & DD.

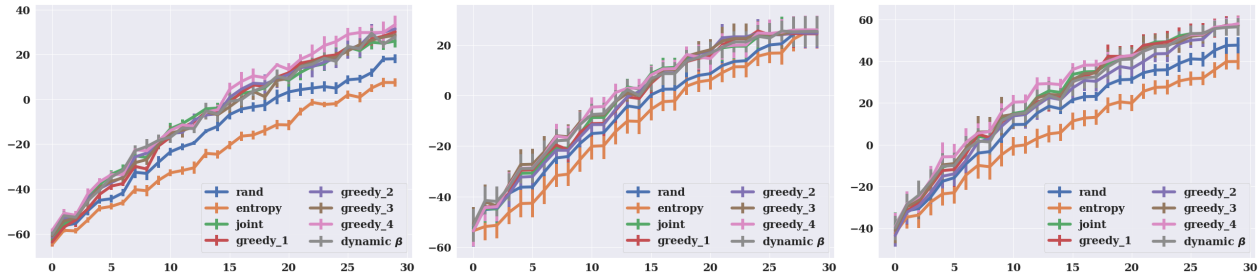

Figure 7: $\sum_i \mathrm{IG}_{i,k}/\beta_i$ vs. iterations k on ODE & SDE.

comparison over the 2-dimensional vector field between agent 1 without collaboration (left) and with collaboration (right). The left plot shows that due to the observational constraints of agent 1 (in terms of restricted time interval), agent 1 can only collect data in a restricted region in the 2-dimensional vector field, which eventually results in the inability to recover the full DE. In contrast, the right plot shows that, as collaboration enables the agent to collect and share data over the entire time interval and hence more coverage of the 2-dimensional vector, the agents can eventually learn to recover the full DE. This is evidenced by the sampled trajectories (orange) following closely to the lines connected by the collected data points (red).

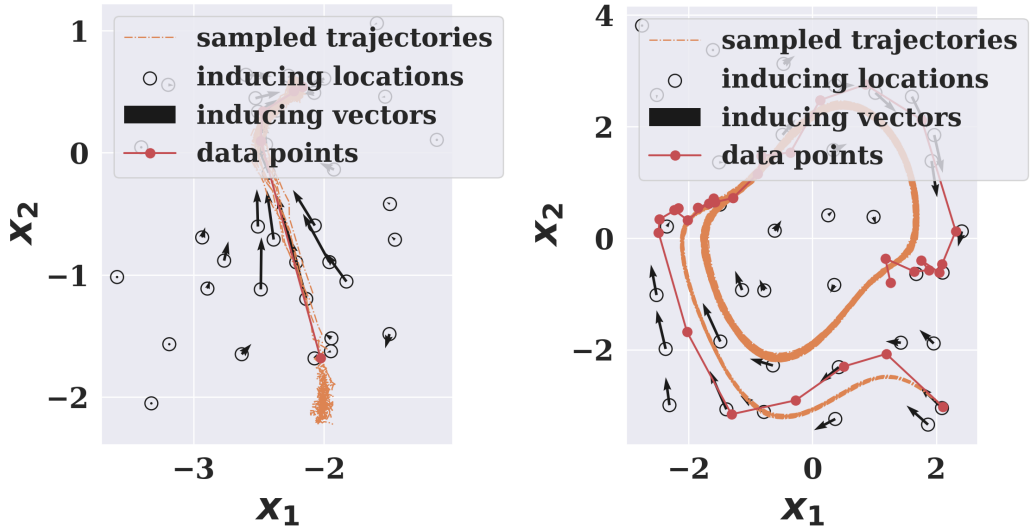

Figure 8: The 2-dimensional vector field in DED-VDP for agent 1. Left (right) corresponds to ind (greedy_1).