

# Bayesian Optimization under **Stochastic** **Delayed Feedback**

Arun Verma\*, Zhongxiang Dai\*, Bryan Kian Hsiang Low

National University of Singapore

ICML 2022



**ICML**  
International Conference  
On Machine Learning

# Motivation



Clinical Trials

Which medicine to try next?



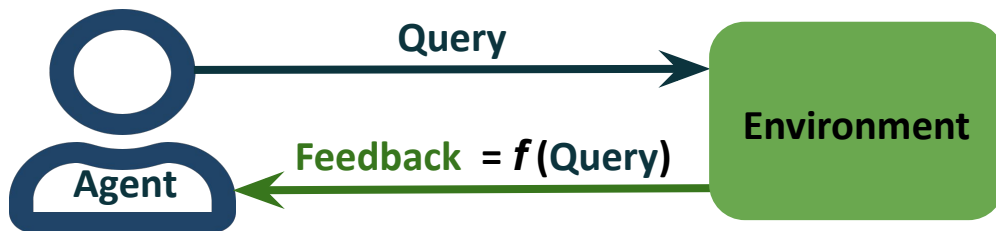
Movie Recommendations

Which movies to show next?

**Similar problems:** Hyperparameter tuning, online advertisement, A/B testing, personalization medicines, and many more.

# Sequential Decision-Making

- In each iteration, agent (or decision-maker) selects the next query.



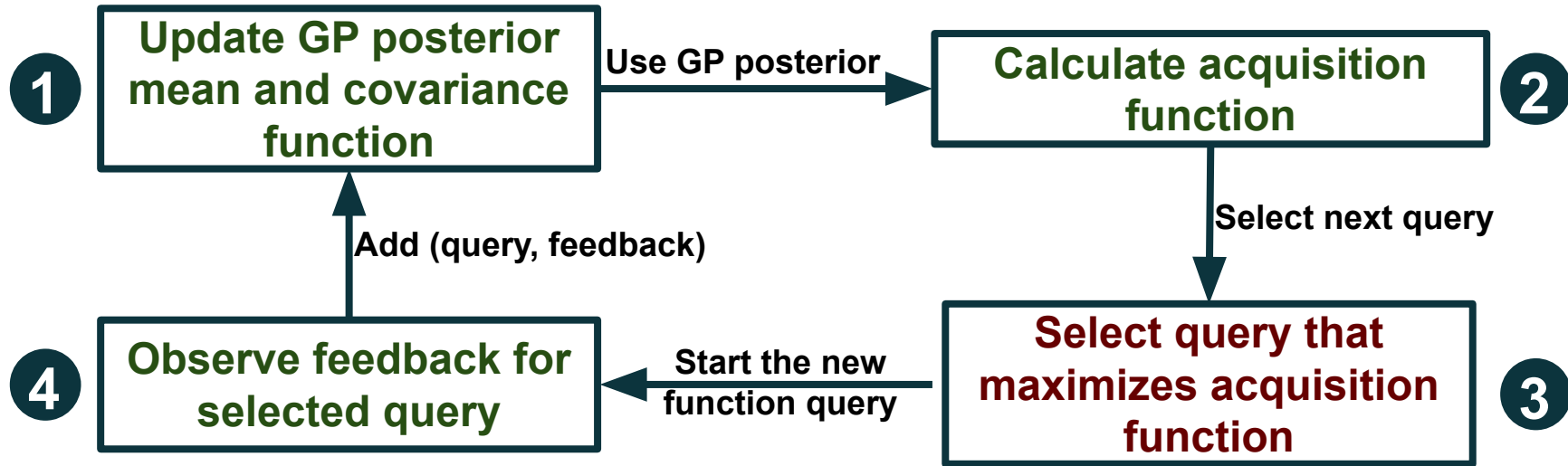
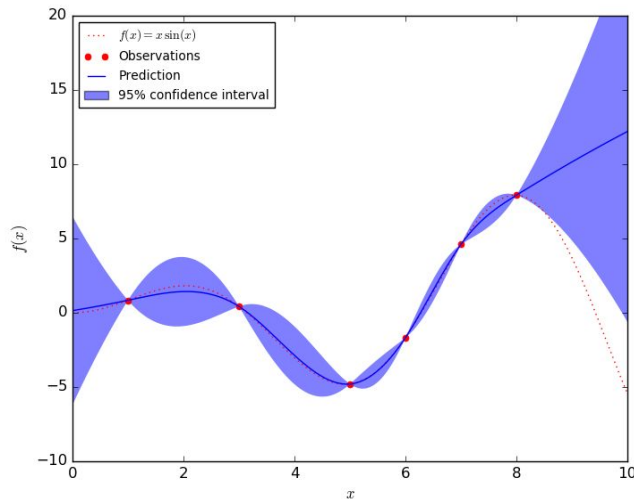
- Environment generates a feedback, which is a function (  $f$  ) of the features of selected query.
- Here,  $f$  is an unknown, non-linear, complex, and blackbox function.

**How to select the next query?**

# Bayesian Optimization (BO)

To choose next function query, BO

- uses a **Gaussian process (GP)** as a surrogate to model the unknown function and
- selects queries by maximizing an **acquisition function** (defined using GP surrogate) that balances **exploration** and **exploitation**.



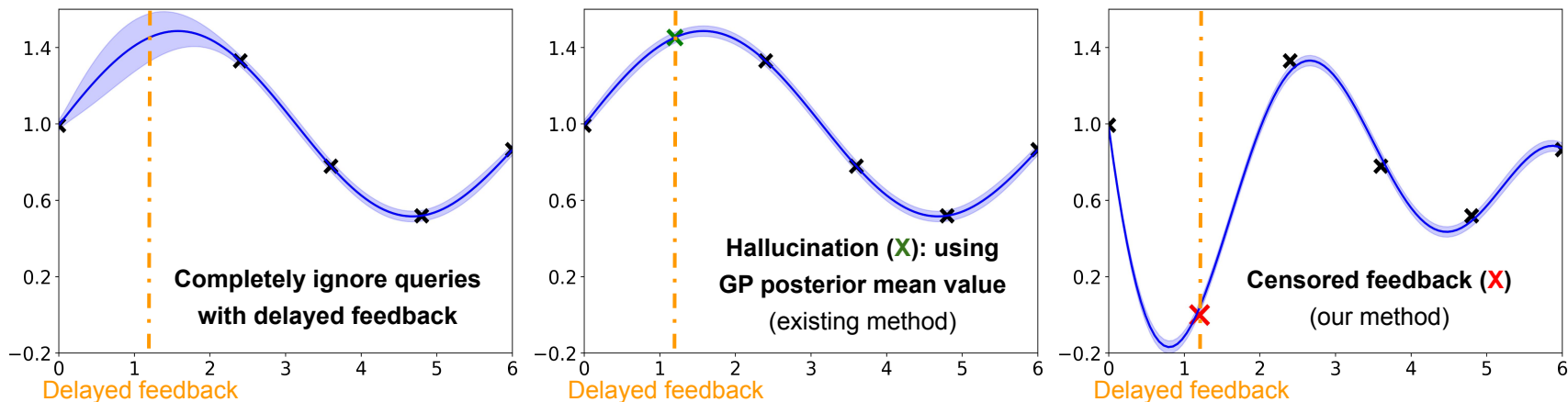
# Stochastic Delayed Feedback

- Feedback is **randomly delayed** in many real-life problems, e.g., hyperparameter tuning, clinical trials, and many more.
- To benefit from the experimental parallelization, the agent must start new function evaluations **without waiting for delayed feedback**.

**How to start a new function query when the observations of the past function queries are randomly delayed?**

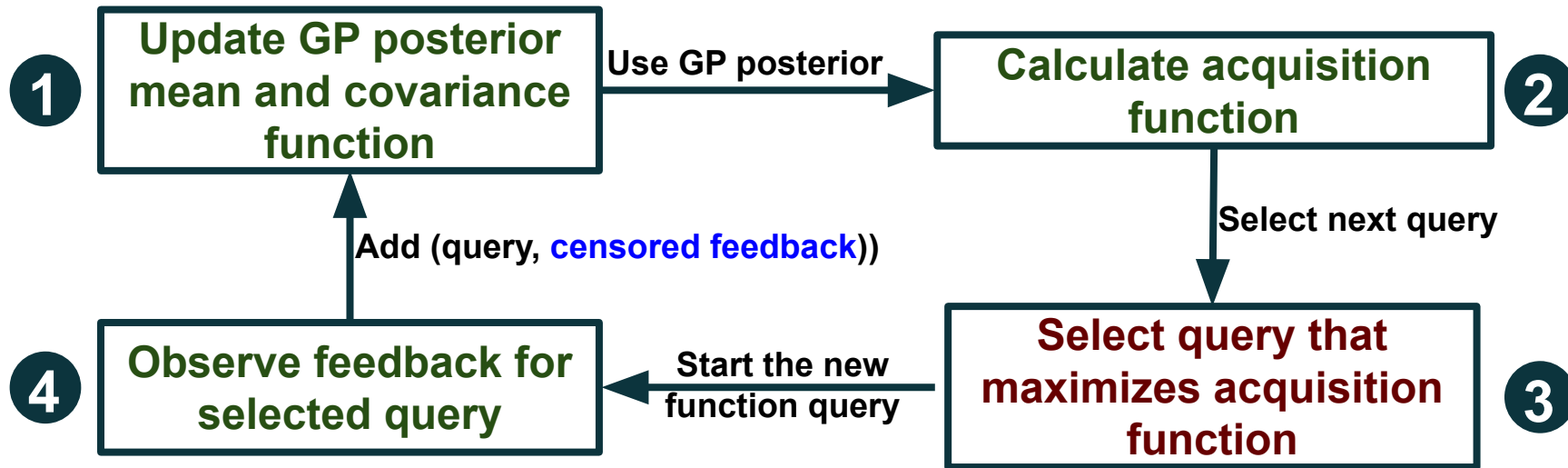
# Censored Feedback

- Delayed feedback is replaced by **censored feedback**, i.e., replacing the unavailable delayed feedback by the **minimum function value**.
- The minimum function value is known in many real-life applications, e.g., the minimum accuracy for hyperparameter tuning of ML models is **0** and a user's minimum response for online recommendation is 'no click' (i.e., **0**).
- Censored feedback (X) can **prevent a selected query from being unnecessarily queried again** by reducing the GP posterior mean around the selected query, hence **better exploration** that leads to better theoretical and empirical performances.



# BO with Censored Feedback

- Replace unavailable delayed feedback by censored feedback in BO.



- With the posterior belief built using censored feedback, we propose **acquisition functions** using upper confidence bound (**UCB**) and Thompson sampling (**TS**).

# Results

- We have shown that our censored feedback-based BO algorithms have **sub-linear regret upper bounds**.

- Regret of **UCB** variant (**GP-UCB-SDF**) for  $T$  queries:

$$\tilde{O} \left( \rho_m^{-1} \left( \gamma_T \sqrt{T} + m \gamma_T \right) \right)$$

Probability of observing delayed feedback within the next  $m$  iterations

Maximum information gain from any set of  $T$  function queries

- Regret of **TS** variant (**GP-TS-SDF**) for  $T$  queries:

$$\tilde{O} \left( \rho_m^{-1} \left( \sqrt{T} \gamma_T (\sqrt{\gamma_T} + 1) + m (\gamma_T + \sqrt{T}) \right) \right)$$

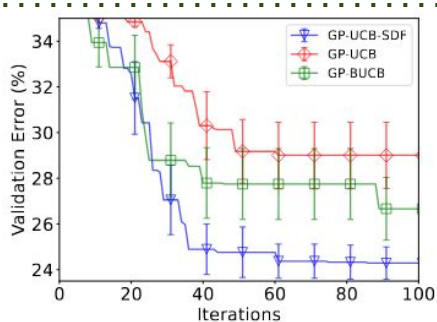
Ignore the logarithmic factors and constants

- **Batch BO** is a special case of our setting and our algorithms have a **better regret bounds** than existing Batch BO algorithms that are based on hallucination.
- We extend our algorithms to **contextual Gaussian process** bandits problem with stochastic delayed feedback

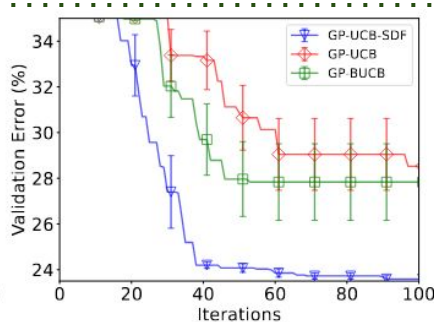


# Experiments

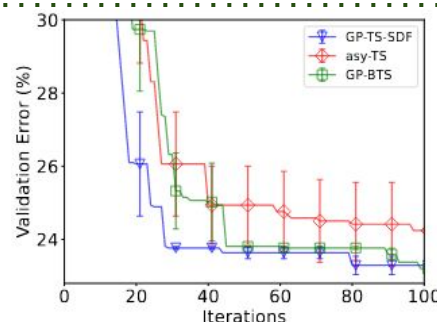
Visit our poster in Poster Session 1.



**Stochastic Delay**

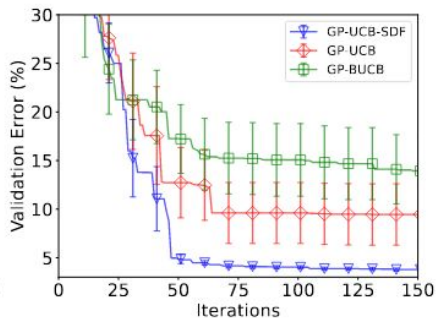


**Deterministic Delay**

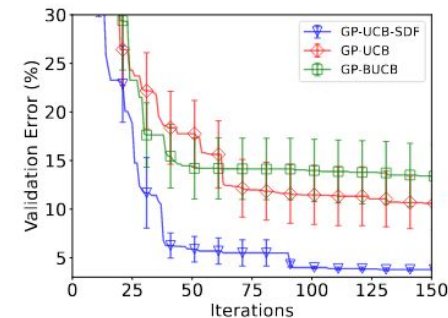


**TS-based Algorithm**

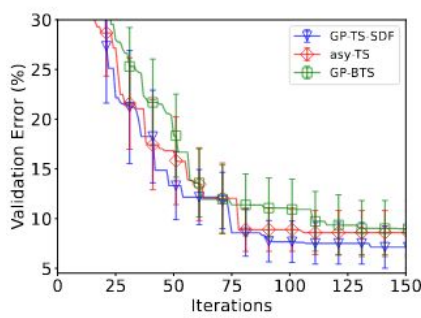
**Hyperparameter tuning of SVM**



**Stochastic Delay**



**Deterministic Delay**



**TS-based Algorithm**

**Hyperparameter tuning of CNN**