Distributed Thompson sampling under constrained communication

Saba Zerefa, Zhaolin Ren, Haitong Ma, Na Li

***Abstract*— In Bayesian optimization, a black-box function is maximized via the use of a surrogate model. We apply distributed Thompson sampling, using a Gaussian process as a surrogate model, to approach the multi-agent Bayesian optimization problem. In our distributed Thompson sampling implementation, each agent receives sampled points from neighbors, where the communication network is encoded in a graph; each agent utilizes their own Gaussian process to model the objective function. We demonstrate theoretical bounds on Bayesian simple regret and Bayesian average regret, where the bound depends on the structure of the communication graph. Unlike in batch Bayesian optimization, this bound is applicable in cases where the communication graph amongst agents is constrained. When compared to sequential single-agent Thompson sampling, our bound guarantees faster convergence with respect to time as long as the communication graph is connected. We confirm the efficacy of our algorithm with numerical simulations on traditional optimization test functions, illustrating the significance of graph connectivity on improving regret convergence.**

***Index Terms*— Bayesian optimization, Thompson sampling, Gaussian process, Distributed optimization, Multi-agent systems**

I. Introduction

Black-box stochastic optimization involves solving problems where the objective function is not explicitly known and can only be accessed through noisy evaluations [1]. These challenges frequently arise in domains where the evaluation process is costly and uncertain, such as hyperparameter tuning in machine learning [2], [3], simulation-based optimization [4], and experimental design [5]. A variety of methods have been developed to tackle these problems, including evolutionary algorithms [6], particle swarm optimization [7], and finite-difference methods [8]. Among these, Bayesian optimization (BO) [9], [10], has emerged as a particularly powerful framework. In contrast to the aforementioned black-box stochastic optimization algorithms which tend to be model-free, by leveraging a probabilistic surrogate model, often a Gaussian process (GP) [11], BO not only handles the stochastic nature of the evaluations but also balances exploration and exploitation given an appropriately chosen surrogate-based sampling strategy. This data-efficient approach makes BO especially well-suited for optimizing expensive, noisy black-box functions [12]. Moreover, theoretically, BO is also known to satisfy finite-time convergence guarantees to global optima (which we note comes at the cost of a dependence on a term that depends on the complexity of the kernel used to model the underlying function) [12]. To the best of our knowledge, apart from BO, finite-time convergence rates in stochastic optimization are only available for finite-difference type methods. However, due to the local nature of finite-difference type methods, the corresponding finite-time convergence rates for such methods only guarantee convergence to stationary points [8], [13], in contrast to the convergence to global optima achieved by BO algorithms.

In BO, the generation of new sampling points is based on the current surrogate model. A good sampling strategy should balance exploration and exploitation of the current surrogate, which is key for efficient optimization. Common sampling strategies include acquisition function-based approaches such as expected improvement (EI) [10] and BO-upper confidence bound (UCB) [14]. Another popular sampling strategy is Thompson sampling, where the next query point is selected as the optimizer of a random function realization sampled from the current posterior [15], [16]. To evaluate algorithm performance, *regret* is studied, which quantifies the gap between the performance of sampled points and the global optimum [17]. Types of regret include simple regret, which measures the gap between the optimal value and the performance of the best queried point [18], and cumulative regret, which measures the sum of the gaps between the optimal value and the performance of each queried point [19], [20].

We are interested in multi-agent BO, where multiple agents can sample the objective function at a single timestep. Much of existing multi-agent BO literature studies batch BO, in which a central coordinator has access to each agent's acquired information [21], [22]. It then computes the sampling decisions for all agents, and communicates these decisions to each agent. These decisions are disseminated in batches, allowing multiple agents to simultaneously sample points, parallelizing the optimization process [23], [24].

Centralized approaches are inapplicable in distributed cases, in which there is no centralized coordinator and each agent must possess a local instance of the optimization algorithm [25]. Additionally, they often do not scale well, as they require a central coordinator to manage the processing of all agents' data. Distributed networks are prevalent in real-world applications, such as in multi-robot source seeking and sensor networks [21], [26]. It may not be the case that all agents have access to all prior sampled points as in the batch setting - communication may be constrained, where some agents are able to communicate with specific other agents [27]. These constraints may be due to limited communication capacity or computational capacity of the agents, or due

*This work was supported by NSF AI institute 2112085, NSF ASCENT, 2328241, and NSF CNS: 2003111.

S. Zerefa, Z. Ren, H. Ma, and N. Li are affiliated with Harvard School of Engineering and Applied Sciences. Emails: szerefa@g.harvard.edu, zhaolinren@g.harvard.edu, haitongma@g.harvard.edu, and nali@seas.harvard.edu.

to physical proximity constraints. Prior literature providing theoretical guarantees for distributed Bayesian optimization require fully connected communication graphs, even in asynchronous cases [22], [28], and thus are inapplicable in settings with constrained communication. In this work, we study the distributed setting with constrained communication, in which at each round, agents send their sampled points to their neighbors and receive points sampled by their neighbors.

Our contribution: We propose a distributed Thompson sampling algorithm for the multi-agent Bayesian optimization problem under constrained communication. In the algorithm, each agent uses their own GP model to pick sampling points via Thompson sampling, and shares the queried points with its neighbors. We then provide provable guarantees for the proposed distributed BO (Corollary 3.2), illustrating a Bayesian simple regret bound of $\tilde{O}\left(\sqrt{\frac{1}{t|V_{max}|}}\right)$, where $|V_{max}|$ is the size of the largest complete subgraph of the communication network G and t is the number of optimization rounds. We note that this convergence speed is $O(\sqrt{|V_{max}|})$ times better than the best known simple regret rate for sequential single-agent BO, which is $\tilde{O}\left(\sqrt{\frac{1}{t}}\right)$ [14]. The analysis also extends to characterizing the Bayesian average regret (Theorem 3.3), which is shown to be $\tilde{O}\left(\frac{\sqrt{n}}{\sqrt{Mt}}\right)$, where M is the number of agents, and n denotes the smallest number L such that the graph G can be decomposed into L disjoint complete subgraphs. This implies then that the average regret bound is smaller for graphs with higher connectivity, which can be decomposed into a few large disjoint complete subgraphs. We numerically test our algorithm on two standard optimization test functions [29] with Erdős-Rényi graphs, demonstrating the efficiency of our algorithm. We find that lower regret is achieved with graphs of higher connectivity, supporting our theoretical results.

II. Problem Formulation and Preliminaries

A. Problem Formulation

For a compact set $\mathcal{X} \subset \mathbb{R}^d$, consider an unknown continuous function $f : \mathcal{X} \to \mathbb{R}$, with optimizer x^*. The goal is to find the maximum of this function, where we are only able to sample f through expensive and noisy evaluations. We assume any of M agents can query f at any point and receive a noisy value $y = f(x) + \epsilon$, with $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. Agents query f throughout a total of T iterations. For agent $i \in \{1, \ldots, M\}$ and iteration $t \in \{1, \ldots, T\}$, $x_{t,i}$ is the query point, and $y_{t,i}$ is the corresponding evaluation. Define $X_{t,i} = \{x_{1,i}, \ldots, x_{t,i}\}$, $Y_{t,i} = \{y_{1,i}, \ldots, y_{t,i}\}$ to be the queries and evaluations made by agent i up to time t. The communication network of M agents is described by graph $G = (V, E)$, where $|V| = M$, and $E \subset \{\{i, j\} : i, j \in V, i \neq j\}$. An unordered pair $\{i, j\} \in E$ if agents i and j are able to communicate with each other. Additionally, we denote the set of neighbors of agent i as $N(i) = \{j : \{i, j\} \in E\}$. The data accessible to agent i at time t is $D_{t,i} = \{(x_{\tau,j}, y_{\tau,j})\}_{j \in N(i) \cup i, \tau < t}$. The set $D_{t,i}$ contains all sampled points up to time t by agent i and its neighbors. We do not make any assumptions regarding the structure of the communication network. The graph may even be unconnected. Our analysis will show how the graph structure affects the algorithm's performance.

B. Gaussian Process

We use a Gaussian process (GP) to model our unknown objective function f in our BO setting. Recall the unknown continuous objective function $f : \mathcal{X} \to \mathbb{R}$. Let $\mathbf{X}_{D_t} = \{x_1, x_2, \ldots, x_t\}$, where x_j is the jth evaluated point. Let $k : \mathcal{X}^2 \to \mathbb{R}$ be a kernel function selected based on prior beliefs about f. Define

$$\mu_{D_t}(x) = \mathbf{k}_t(x)^\intercal (\mathbf{K}_{D_t} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{y}_{D_t}$$
$$k_{D_t}(x, x') = k(x, x') - \mathbf{k}_{D_t}(x)^\intercal (\mathbf{K}_{D_t} + \sigma_n^2 \mathbf{I})^{-1} \mathbf{k}_{D_t}(x'),$$

where $\mathbf{K}_{D_t} := [k(x', x'')]_{x', x'' \in \mathbf{X}_{D_t}}$, $\mathbf{k}_{D_t}(x) := [k(x', x)]_{x' \in \mathbf{X}_{D_t}}$ and $\mathbf{y}_{D_t} = \{f(x') + \epsilon'\}_{x' \in \mathbf{X}_{D_t}}$, where $\epsilon' \sim \mathcal{N}(0, \sigma_\epsilon^2)$. Thus we can define our GP, in which we denote $f | \mathcal{F}_{D_t} \sim GP(\mu_{D_t}(x), k_{D_t}(x, x'))$. Note that due to the nature of the GP, it is the case that for any $x \in \mathcal{X}$, $f(x) | \mathcal{F}_{D_t} \sim N(\mu_{D_t}(x), \sigma^2_{D_t}(x))$, where $\sigma^2_{D_t}(x) = k_{D_t}(x, x)$ [11]. Furthermore, recall the distributed multi-agent setting, where each of M agents have access to queried points in set $D_{t,i}$, where $D_{t,i}$ and $D_{t,j}$ may not be equal for distinct agents i and j. In our distributed setting, each agent i has a unique GP model of f at time t, $\mathcal{GP}_{t,i}$, since the data $D_{t,i}$ available to each agent i is different. Thus we denote $f \mid \mathcal{F}_{D_{t,i}} \sim \mathcal{GP}_{t,i}(\mu_{D_{t,i}}(x), k_{D_{t,i}}(x, x'))$.

C. Regret

Our metric for algorithm performance is regret, which is an assessment of the quality of sampled points. We consider the simple regret, which is the difference between the optimal value of the function and the best value achieved amongst the previous queried points. This definition of regret is useful because optimization settings focus on locating the extrema of a function, and the simple regret tracks the smallest gap between the value at a sampled point and the optimal value. To account for randomness of f in our regret expression, we take expectation of simple regret to yield the following expression, which is called Bayesian simple regret:

$$R_{SB}(t) = \min_{i \in \{1,2,\ldots,M\}, \tau \in \{1,2,\ldots,t\}} \mathbb{E}[f(x^*) - f(x_{\tau,i})] \quad (1)$$

We also consider average regret, which quantifies the difference between the optimal value of the function and the queried value for each sampled point. In average regret, this difference is accumulated across all agents and timesteps, and then averaged by the amount of sampled points. We also take the expectation of this quantity to yield Bayesian average regret:

$$R_{AB}(t) = \frac{1}{tM} \sum_{\tau=1}^{t} \sum_{i=1}^{M} \mathbb{E}[f(x^*) - f(x_{\tau,i})] \tag{2}$$

In our theoretical analysis, we provide bounds on Bayesian simple regret and Bayesian average regret.

D. Thompson Sampling

Thompson sampling is an algorithm for sequential decision making that can be utilized in this context for determining the next point of the objective function to query [15]. When using Thompson sampling in our Bayesian optimization framework, an acquisition function is sampled from the posterior distribution of the Gaussian process. The maximizer of this function is the next query point at which the black-box objective function is sampled. The Gaussian process is then updated with new information from this sample, and the process repeats for the duration of the experiment.

In sequential single-agent Thompson sampling, each subsequent query point is determined based on a single model updated on all prior sampled points. Alternatively in batch Thompson sampling, multiple query points are determined as a set at each round, and the objective function is sampled in parallel [22], [24]. Batch Thompson sampling is advantageous in systems capable of parallelizing, e.g. multi-agent systems, because it allows for convergence in fewer number of rounds than sequential single-agent Thompson sampling.

Batch Thompson sampling is centralized, with all agents having access to the same information. However, this may not be realistic in real-world situations, where communication between agents may be constrained due to bandwidth limitations, computational constrictions, or privacy concerns. In these cases, agents may only have access to the sampled points by few other agents, and thus datasets available to distinct agents may differ. We propose a Thompson sampling algorithm for this constrained communication case, and provide theoretical guarantees for the algorithm.

III. Algorithm: Distributed Thompson Sampling

In our implementation of distributed Thompson sampling, each of M agents have distinct Gaussian processes $\mathcal{GP}_i$ for modeling the objective function. At each time step t, all agents update their GPs with the data history available to them. The agent then queries the objective function at $x_{t,i}$, which is the maximizer of the acquisition function sampled from the posterior GP, $\hat{f}_{t,i} \sim \mathcal{GP}_{t,i}$. Each agent then communicates its sampled point to its neighbors, receives the points sampled by their neighbors, and updates their data history accordingly. The collection of data received by neighbors of agent i at time t is denoted as $C_{t,i} = \{(x_{t,j}, y_{t,j})\}_{j \in N(i)}$. Our method is shown in Algorithm 1. We stress that while we do assume a synchronous global clock, there exists no centralized coordinator in our algorithm that coordinates the queries of the different agents.

Algorithm 1 Distributed Thompson Sampling

```
Place GP prior on f
for i= 1, ..., M do
    Initial data D_{1,i}
    GP_{0,i} ← GP
end for
for t = 1, ..., T do
    for i= 1, ..., M do
        Update posterior GP_{t,i} conditioned on D_{t,i}
        Sample f̂_{t,i} ~ GP_{t,i}
        Choose next query point
            x_{t,i} ← arg max_x f̂_{t,i}(x)
        Observe y_{t,i}
        Broadcast (x_{t,i}, y_{t,i}) to neighbors N(i);
        Collect evaluations C_{t,i} from neighbors N(i)
        Update data history D_{t+1,i} ← D_{t,i} ∪ C_{t,i} ∪
    {(x_{t,i}, y_{t,i})}
    end for
end for
```

A. Theoretical Result

We analyze the performance of the distributed Thompson sampling algorithm on the Bayesian simple regret and Bayesian average regret metrics. Our regret bound depends on the number of timesteps T and the structure of the agent communication graph G. As in prior work, we utilize notions from information theory in our regret bound [30].

Our regret bound involves the Maximum Information Gain (MIG), which is a constant that captures the difficulty of the objective function. MIG has been shown to be bounded for several kernel functions commonly used with GPs, including Squared Exponential and Matérn kernels, the latter of which was used in our numerical implementation [14].

Let $D = \{x_1, \ldots, x_t\} \subset \mathcal{X}$, and define $y_D = \{(x, f(x) + \epsilon) : x \in D\}$. The MIG is denoted as

$$\Psi_t = \max_{D \subset \mathcal{X}, |D|=t} I(f; y_D), \tag{3}$$

where I is the Shannon Mutual Information. The MIG Ψ_t represents the largest mutual information gain from f by sampling t points. Additionally, for any positive integer n, we define the constant ξ_n, which bounds the information gain of the current round of evaluations [22]. Suppose $|D| = t$ points were already sampled, and i points are being queried in the current round of evaluations, with $i < n$; denote these points in set A, where $A \subset \mathcal{X}$, and $y_A = \{(x, f(x) + \epsilon) : x \in A\}$. Then for $i \geq 1$, ξ_n satisfies

$$\max_{A \subset \mathcal{X}, |A|<n} I(f; y_A | y_D) \leq \frac{1}{2} \log(\xi_n). \tag{4}$$

In the following theorem, we provide a bound for Bayesian simple regret for an M agent system with communication graph G. After proving Theorem 3.1, we will introduce and discuss its implication, Corollary 3.2, which is our main convergence result.

Theorem 3.1. *Suppose $k(x, x') \leq 1$ for all x, x'. Let $G_s = (V_s, E_s)$ be a complete subgraph of G. Then the Bayesian simple regret after t timesteps satisfies $R_{SB}(t) \leq \frac{C_1}{t|V_s|} + \sqrt{\frac{C_2 \xi_{|V_s|} \beta_t \Psi_{t|V_s|}}{t|V_s|}}$, where $\beta_t = 2\log(t^2|V_s||\mathcal{X}|)$, $C_1 = \frac{\sqrt{2}\pi^{3/2}}{12}$, and $C_2 = \frac{2}{\log(1+\sigma_\epsilon^{-2})}$.*

Proof. The structure of our proof follows techniques from Kandasamy et al. [22]. We aim to provide a bound on Bayesian simple regret. Our proof begins by noting that we can bound Bayesian simple regret by the average expected regret of any subset of agents. We then decompose this into three sums, each of which utilize a confidence function $U_{t,i}(\cdot)$. We bound each of these sums by incorporating notions from information theory, which allows us to use information gain constants introduced in Equations 3 and 4 to analyze the efficacy of the sampling process.

Recall from equation 1 that $R_{SB}(t) = \min_{i\in\{1,2,\ldots,M\},\tau\in\{1,2,\ldots,t\}} \mathbb{E}[f(x^*) - f(x_{\tau,i})]$. Define $U_{t,i}(x) = \mu_{D_{t,i}}(x) + \beta_t^{1/2}\sigma_{D_{t,i}}(x)$. Note that, since $R_{SB}(t)$ represents a minimum value of the expression $\mathbb{E}[f(x^*) - f(x_{\tau,i})]$, it is upper bounded by the average of this expression across any subset of agents. Enumerate the agents in the complete graph $G_s = (V_s, E_s)$ as $\{1, 2, \ldots, |V_s|\}$. We reindex the M agents such that agents $i \in V_s$ are indexed $\{1, 2, \ldots, |V_s|\}$ and agents $j \in V\backslash V_s$ are indexed $\{|V_s| + 1, \ldots, M\}$. Thus, we may write

$$R_{SB}(t) \leq \frac{1}{t|V_s|}\sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}[f(x^*) - f(x_{\tau,i})].$$

We will then focus on upper bounding the expression $\sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}[f(x^*) - f(x_{\tau,i})]$.

$$\begin{aligned}
&\sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}[f(x^*) - f(x_{\tau,i})] \\
&= \sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}[f(x^*) - U_{\tau,i}(x^*) + U_{\tau,i}(x^*) - U_{\tau,i}(x_{\tau,i}) \\
&\qquad + U_{\tau,i}(x_{\tau,i}) - f(x_{\tau,i})] \\
&= \sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\underbrace{\mathbb{E}[f(x^*) - U_{\tau,i}(x^*)]}_{S1} + \underbrace{\mathbb{E}[U_{\tau,i}(x^*) - U_{\tau,i}(x_{\tau,i})]}_{S2} \\
&\qquad + \underbrace{\mathbb{E}[U_{\tau,i}(x_{\tau,i}) - f(x_{\tau,i})]}_{S3}
\end{aligned}$$

We will now bound each of these sums.

S1. Let's begin by upper bounding the sum $S1 = \sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}[f(x^*) - U_{\tau,i}(x^*)]$.

$$S1 = \sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}[f(x^*) - U_{\tau,i}(x^*)] \tag{5}$$

$$\leq \sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}\big[\mathbb{E}[\mathbb{I}\{f(x^*) > U_{\tau,i}(x^*)\}(f(x^*) - U_{\tau,i}(x^*))\,|\,\mathcal{F}_{D_{t,i}}]\big] \tag{6}$$

$$\leq \sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\sum_{x\in\mathcal{X}}\frac{e^{-\beta_\tau/2}}{\sqrt{2\pi}} \leq \sum_{\tau=1}^{t}\frac{1}{\tau^2\sqrt{2\pi}} \leq \frac{\sqrt{2}\pi^{3/2}}{12} \tag{7}$$

Line (6) upper bounds $S1$ by positive terms by making use of the indicator function $\mathbb{I}\{f(x^*) > U_{\tau,i}(x^*)\}$, which takes the value of 1 when the condition $f(x^*) > U_{\tau,i}(x^*)$ is satisfied and 0 otherwise. Line (7) utilizes Lemma I.1 in Appendix I of our supplement [31] to bound the expectation of positive terms in a normal distribution. Additionally, Line (7) results from substituting for β_τ, and the fact that $\sum_{j=1}^{\infty}\frac{1}{j^2} = \frac{\pi^2}{6}$. Therefore, we have established an upper bound for $S1$.

S2. We evaluate the expression $S2 = \sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}[U_{\tau,i}(x^*) - U_{\tau,i}(x_{\tau,i})]$. First let's focus on the interior of the summation, $\mathbb{E}[U_{\tau,i}(x^*) - U_{\tau,i}(x_{\tau,i})]$. We will proceed to show that this expression evaluates to 0.

By the law of total expectation, $\mathbb{E}[U_{\tau,i}(x^*) - U_{\tau,i}(x_{\tau,i})] = \mathbb{E}[\mathbb{E}[U_{\tau,i}(x^*) - U_{\tau,i}(x_{\tau,i})]|D_{t,i}]$. Because $x_{t,i}$ is sampled from the posterior distribution of $x^*|D_{t,i}$, we must have that $x_{t,i}$ and x^* have the same distribution after conditioning on the acquired data, and thus $x_{t,i}|D_{t,i} \sim x^*|D_{t,i}$. We also notice that $U_{t,i}$ is deterministic when conditioned on $D_{t,i}$. Thus we have that $\mathbb{E}[\mathbb{E}[U_{\tau,i}(x^*) - U_{\tau,i}(x_{\tau,i})]|D_{t,i}]] = 0$, and consequently $S2 = 0$.

S3. Lastly, we bound $S3 = \sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}[U_{\tau,i}(x_{\tau,i}) - f(x_{\tau,i})]$. For the evaluation of this sum, we will introduce some additional notation. Denote $\bar{D}_t = \{(x_{\tau,j}, y_{\tau,j})\}_{j\in V_s, \tau<t}$. We can think of $\bar{D}_t$ as representing the data acquired by round t in a batch setting exclusively by agents in V_s. Additionally, define $\sigma_{t,i}(x) := \sigma(x) \mid \bar{D}_t \cup \{(x_{t,j}, y_{t,j})\}_{j\in V_s, j<i}$. We pick to define $\sigma_{t,i}(\cdot)$ in this way to impose an ordering of the acquired data; in this sense, $\sigma_{t,i}(\cdot)$ depends on $(t-1)|V_s| + i - 1$ previously sampled points. With this additional notation, we are well-equipped to upper bound $S3$.

$$S3 = \sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}[U_{\tau,i}(x_{\tau,i}) - f(x_{\tau,i})] \tag{8}$$

$$= \sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}[\mathbb{E}[\mu_{D_{\tau,i}}(x_{\tau,i}) + \beta_\tau^{1/2}\sigma_{D_{\tau,i}}(x_{\tau,i}) - f(x_{\tau,i})]|D_{\tau,i}]] \tag{9}$$

$$= \sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}[\beta_\tau^{1/2}\sigma_{D_{\tau,i}}(x_{\tau,i})] \tag{10}$$

$$\leq \beta_t^{1/2}\sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}[\sigma_{D_{\tau,i}}(x_{\tau,i})] \tag{11}$$

$$\leq \beta_t^{1/2}\sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}[\sigma_{\bar{D}_\tau}(x_{\tau,i})] \tag{12}$$

$$\leq \beta_t^{1/2}\mathbb{E}\left[\sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\sigma_{\tau,i}(x_{\tau,i})\exp(I(f;\{y_{\tau,j}\}_{j<i}|y_{\bar{D}_\tau})\right] \tag{13}$$

$$\leq \beta_t^{1/2}\mathbb{E}\left[\sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\sigma_{\tau,i}(x_{\tau,i})\xi_i^{1/2}\right] \tag{14}$$

$$\leq \beta_t^{1/2}\xi_{|V_s|}^{1/2}\mathbb{E}\left[\sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\sigma_{\tau,i}(x_{\tau,i})\right] \tag{15}$$

$$\leq \beta_t^{1/2}\xi_{|V_s|}^{1/2}\mathbb{E}\left[\left(t|V_s|\sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\sigma_{\tau,i}^2(x_{\tau,i})\right)^{1/2}\right] \tag{16}$$

$$\leq \beta_t^{1/2}\xi_{|V_s|}^{1/2}\sqrt{\frac{2t|V_s|\Psi_{t|V_s|}}{\log(1+\sigma_\epsilon^{-2})}} \leq \sqrt{\frac{2\xi_{|V_s|}t|V_s|\beta_t\Psi_{t|V_s|}}{\log(1+\sigma_\epsilon^{-2})}} \tag{17}$$

Lines (9) and (10) follow by the law of total expectation. Line (11) follows by noting that β_τ is increasing with τ. Recall that because G_s is a subgraph of G, for all agents $i \in V_s$, $V_s \subset \{N(i) \cup i\}$. Therefore, $\bar{D}_\tau$ is contained in $D_{\tau,i}$ for all agents $i \in V_s$. In essence, $\bar{D}_\tau \subset D_{\tau,i} \implies \sigma_{\bar{D}_\tau} \geq \sigma_{D_{\tau,i}}$. Line (12) applies this property. Line (13) follows from Lemma I.4 in Appendix I in our supplement [31], which introduces ξ_i to bound the information gain of the current set of evaluations. Line (14) follows from the definition of ξ_i, which was stated in Equation 4. Line (15) is a consequence of the fact that ξ_i is increasing with i. Line (16) follows from application of the Cauchy-Schwarz inequality, and line (17) is a consequence of Lemma I.3, which bounds the sums of the posterior variances by the MIG term.

Recall that $\sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}[f(x^*) - f(x_{\tau,i})] = S1 + S2 + S3$. Thus, we have shown that $\sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}[f(x^*) - f(x_{\tau,i})] \leq \frac{\sqrt{2}\pi^{3/2}}{12} + \sqrt{\frac{2\xi_{|V_s|}t|V_s|\beta_t\Psi_{t|V_s|}}{\log(1+\sigma_\epsilon^{-2})}}$. However, we wish to bound $R_{SB}(t)$. Since $R_{SB}(t) \leq \frac{1}{t|V_s|}\sum_{\tau=1}^{t}\sum_{i=1}^{|V_s|}\mathbb{E}[f(x^*) - f(x_{\tau,i})]$, it directly follows that $R_{SB}(t) \leq \frac{\sqrt{2}\pi^{3/2}}{12t|V_s|} + \sqrt{\frac{2\xi_{|V_s|}\beta_t\Psi_{t|V_s|}}{t|V_s|\log(1+\sigma_\epsilon^{-2})}}$, thus concluding our proof. □

Picking G_s to be the largest complete subgraph of the communication network G yields the following corollary.

Corollary 3.2. *Suppose $k(x,x') \leq 1$ for all x, x'. Let $G_{max} = (V_{max}, E_{max})$ be the largest complete subgraph of G. Then the Bayesian simple regret after t timesteps satisfies $R_{SB}(t) \leq \frac{C_1}{t|V_{max}|} + \sqrt{\frac{C_2\xi_{|V_{max}|}\beta_t\Psi_{t|V_{max}|}}{t|V_{max}|}}$, where $\beta_t = 2\log(t^2|V_{max}||\mathcal{X}|)$, $C_1 = \frac{\sqrt{2}\pi^{3/2}}{12}$, and $C_2 = \frac{2}{\log(1+\sigma_\epsilon^{-2})}$.*

From the corollary, we see that $R_{SB}(t) = \tilde{O}\left(\sqrt{\frac{\xi_{|V_{max}|}\Psi_{t|V_{max}|}}{t|V_{max}|}}\right)$. We note that the term $\Psi_{t|V_{max}|}$ corresponds to the maximal mutual information gain from $t|V_{max}|$ observations, and that this quantity depends only logarithmically on $t|V_{max}|$ for standard kernels such as the squared exponential kernel. For more details, we refer to Appendix II in our supplement [31]. We note that the term $\xi_{|V_{max}|}$ is the price we pay for the delay in communication within the subgraph G_{max}, and is a standard term that arises in multi-agent Bayesian optimization. By an appropriately designed initialization phase, this term can actually be reduced to $\tilde{O}(1)$ (see Appendix B.3 in [24] for more details). Thus, compared to the sequential single-agent case with t rounds which has simple regret $\tilde{O}\left(\sqrt{\frac{1}{t}}\right)$ [14], our algorithm satisfies a regret of $\tilde{O}\left(\sqrt{\frac{1}{t|V_{max}|}}\right)$, i.e. an improvement of $\sqrt{\frac{1}{|V_{max}|}}$, demonstrating the benefit of the network structure. We also introduce the following result on Bayesian average regret.

Theorem 3.3. *Suppose $k(x,x') \leq 1$ for all x, x'. Let $\{G_k\}_{k\in\{1,\dots,n\}}$ be a collection of n disjoint complete subgraphs of communication graph $G = (V,E)$, where $G_k = (V_k, E_k)$, and $\cup_{k\in\{1,\dots,n\}}V_k = V$. Then the Bayesian average regret after t timesteps satisfies $R_{AB}(t) \leq \frac{1}{M}\sum_{k=1}^{n}|V_k|(\frac{C_1}{t|V_k|} + \sqrt{\frac{C_2\xi_{|V_k|}\beta_t\Psi_{t|V_k|}}{t|V_k|}})$, where $\beta_t = 2\log(t^2M|\mathcal{X}|)$, $C_1 = \frac{\sqrt{2}\pi^{3/2}}{12}$, and $C_2 = \frac{2}{\log(1+\sigma_\epsilon^{-2})}$.*

Proof. The structure of our proof is similar to that of Theorem 3.1. We aim to provide a bound on the Bayesian average regret affiliated with agents in communication graph G by bounding the Bayesian average regret affiliated with complete subgraphs of graph G. Let $G = (V,E)$ be the communication graph for the M agents. We can construct a collection of n disjoint complete subgraphs $\{G_k\}_{k\in\{1,\dots,n\}}$, where each $G_k = (V_k, E_k)$ is a subgraph of G, with $\cup_{k\in\{1,2,\dots,n\}}V_k = V$.

Recall from Equation 2 that Bayesian average regret $R_{AB}(t) = \frac{1}{tM}\sum_{\tau=1}^{t}\sum_{i=1}^{M}\mathbb{E}[f(x^*) - f(x_{\tau,i})]$. Recalling the partition of the vertex set V into $\{V_k\}_{k\in\{1,\dots,n\}}$, we may rewrite Bayesian average regret as follows:

$$\begin{aligned}R_{AB}(t) &= \frac{1}{tM}\sum_{\tau=1}^{t}\sum_{i=1}^{M}\mathbb{E}[f(x^*) - f(x_{\tau,i})] \\ &= \frac{1}{tM}\sum_{\tau=1}^{t}\sum_{k=1}^{n}\sum_{i\in V_k}\mathbb{E}[f(x^*) - f(x_{\tau,i})]\end{aligned}$$

As an intermediary step in proving Theorem 3.1, recall that for a complete graph G_k, we have shown that $\sum_{\tau=1}^{t}\sum_{i\in V_k}\mathbb{E}[f(x^*) - f(x_{\tau,i})] \leq C_1 + \sqrt{C_2\xi_{|V_k|}\beta_t\Psi_{t|V_k|}t|V_k|}$, where $C_1 = \frac{\sqrt{2}\pi^{3/2}}{12}$, $C_2 = \frac{2}{\log(1+\sigma_\epsilon^{-2})}$, and $\beta_t = 2\log(t^2|V_k||\mathcal{X}|)$. Since $|V_k| \leq M$, this inequality also holds for $\beta_t = 2\log(t^2M|\mathcal{X}|)$.

Equipped with this bound affiliated with G_k, we revisit our Bayesian average regret expression:

$$\begin{aligned}R_{AB}(t) &= \frac{1}{tM}\sum_{\tau=1}^{t}\sum_{k=1}^{n}\sum_{i\in V_k}\mathbb{E}[f(x^*) - f(x_{\tau,i})] \\ &= \frac{1}{tM}\sum_{k=1}^{n}\sum_{\tau=1}^{t}\sum_{i\in V_k}\mathbb{E}[f(x^*) - f(x_{\tau,i})] \\ &\leq \frac{1}{tM}\sum_{k=1}^{n}C_1 + \sqrt{C_2\xi_{|V_k|}\beta_t\Psi_{t|V_k|}t|V_k|} \\ &\leq \frac{1}{M}\sum_{k=1}^{n}\left(\frac{C_1}{t} + \sqrt{\frac{C_2\xi_{|V_k|}\beta_t\Psi_{t|V_k|}|V_k|}{t}}\right)\end{aligned}$$

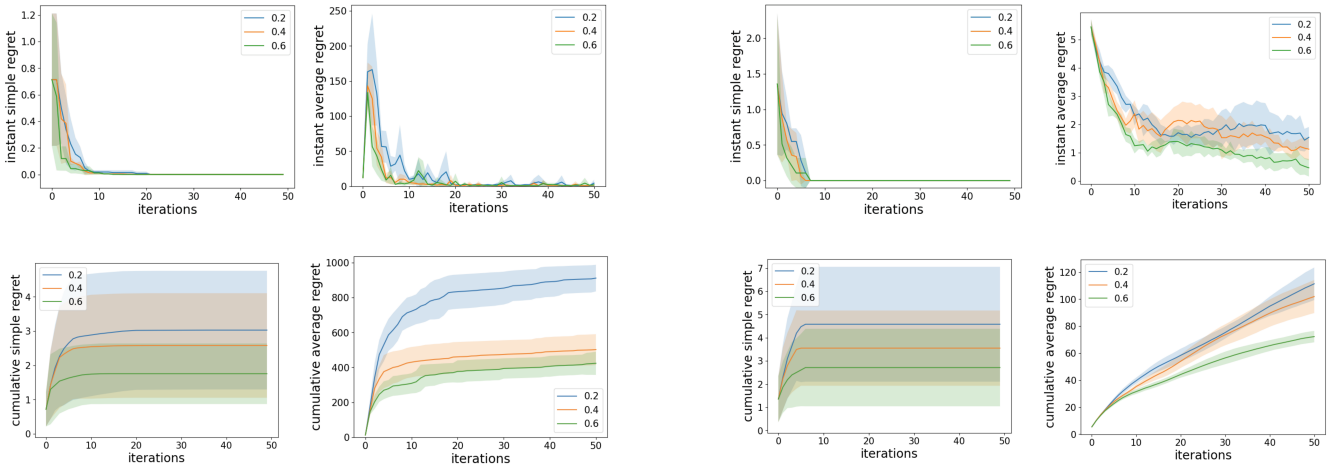


Fig. 1: Regret analysis of numerical simulations with 20 agents on Erdős-Rényi random graphs with connectivity probability 0.2 (blue), 0.4 (orange), and 0.6 (green), on Rosenbrock (left) and Ackley (right) objective functions.

Thus, we have developed an upper bound for Bayesian average regret, concluding our proof. □

This result, similar to that of Theorem 3.1, follows from considering the cumulative regret of complete subgraphs of communication graph G. As a simple corollary, by applying Cauchy-Schwarz to bound the term $\sum_{i=1}^{n}\sqrt{|V_i|} \leq \sqrt{n}\sqrt{\sum_{i=1}^{n}|V_i|} = \sqrt{Mn}$, we find that $R_{AB}(t) = \tilde{O}\left(\frac{\sqrt{n}}{\sqrt{Mt}}\right)$, where n is the smallest number L such which G can be decomposed into L disjoint complete subgraphs.

IV. Numerical Experiments

A. Simulation

In the numerical implementation, performance was assessed utilizing the following regret metrics.

We define the *Instant simple regret* R_B, and its sum $\overline{R_B}$, as follows:

$$R_B(t) = f^* - \max_{\substack{i\in\{1,2,\dots M\},\\ \tau\in\{1,2,\dots,t\}}} f(x_{\tau,i}), \quad \overline{R_B}(t) = \sum_{\tau=1}^{t} R_B(\tau),$$

where $f^* = \max_{x\in\mathcal{X}} f(x)$. This can be regarded as a numerical proxy for the Bayesian simple regret.

We also define the *Instant average regret* R_A, and its sum, $\overline{R_A}$, as follows:

$$R_A(t) = \frac{1}{M}\sum_{i=1}^{M}\left(f^* - f(x_{\tau,i})\right), \quad \overline{R_A}(t) = \sum_{\tau=1}^{t} R_A(\tau).$$

Numerical results were constructed in a Python implementation built upon the BOTorch package [32]. The code used to generate the simulation and corresponding Figure 1 is available at https://github.com/sabzer/distributed-bo. The Gaussian processes utilized the Matérn kernel with parameter $\nu = \frac{5}{2}$. The numerical simulations were run over $T = 50$ timesteps. Simulations were run based on two test functions for the objective function: Ackley, which has many local maxima and one global minima in the origin, and Rosenbrock, which contains a large valley in which the global minima is situated. The equations of the aforementioned objective functions and their plots are available in Appendix III in our supplement [31]. Since we were solving a maximization problem, we multiplied the canonical definitions of these functions by -1 for the purpose of our simulation. For the communication networks in our simulations, we used Erdős-Rényi random graphs of 20 agents with connectivities of 0.2, 0.4, and 0.6 [33]. The connection probabilities are the probability that each edge from the complete graph of 20 agents appears in the corresponding random graph.

B. Discussion

Our theoretical result bounds Bayesian simple regret, $R_{SB}(t)$, and Bayesian average regret, $R_{AB}(t)$ with the bound dependent on the structure of the communication network between agents. Our distributed Thompson sampling algorithm was able to achieve the extrema of the Ackley and Rosenbrock objective functions in numerical implementation, and thus is effective at the Bayesian optimization task. Our theoretical results suggests lower regret when the largest complete subgraph of the communication graph has a larger number of agents. Our numerical results support this intuition, for in Figure 1, we see better regret convergence for Erdős-Rényi graphs of higher connectivity. This result holds for both Ackley and Rosenbrock objective functions, and for both instant simple regret and instant average regret.

V. Conclusion

In this paper, we proposed a distributed Thompson sampling algorithm to address the multi-agent Bayesian optimization problem under constrained communication. We develop a bound on Bayesian simple regret for this approach, where the bound is dependent on properties of the largest complete subgraph of the graph encoding communication

structure between agents. With our bound, we show that in connected multi-agent communication networks, Bayesian simple regret will converge faster with distributed Thompson sampling than in the sequential single-agent case, with the same number of rounds. Additionally, we demonstrate the efficacy of our algorithm with regret analysis on optimization test functions, illustrating faster convergence with graphs possessing larger connected subgraphs. Future work will focus on developing a tighter regret bound, and further tailoring the distributed Thompson sampling algorithm towards the constrained communication case by leveraging the data communicated between agents.

REFERENCES

[1] J. C. Spall, *Introduction to stochastic search and optimization: estimation, simulation, and control*. John Wiley & Sons, 2005.

[2] J. Bergstra, D. Yamins, and D. Cox, "Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures," in *Proceedings of the 30th International Conference on Machine Learning*, vol. 28, pp. 115–123, PMLR, 2013.

[3] M. Lindauer, K. Eggensperger, M. Feurer, A. Biedenkapp, D. Deng, C. Benjamins, T. Ruhkopf, R. Sass, and F. Hutter, "Smac3: A versatile bayesian optimization package for hyperparameter optimization," *Journal of Machine Learning Research*, vol. 23, no. 54, pp. 1–9, 2022.

[4] R. Y. Rubinstein and D. P. Kroese, *Simulation and the Monte Carlo method*. John Wiley & Sons, 2016.

[5] R.-R. Griffiths and J. M. Hernández-Lobato, "Constrained bayesian optimization for automatic chemical design using variational autoencoders," *Chem. Sci.*, vol. 11, pp. 577–586, 2020.

[6] A. Slowik and H. Kwasnicka, "Evolutionary algorithms and their applications to engineering problems," *Neural Computing and Applications*, vol. 32, pp. 12363–12379, 2020.

[7] F. Marini and B. Walczak, "Particle swarm optimization (pso). a tutorial," *Chemometrics and Intelligent Laboratory Systems*, vol. 149, pp. 153–165, 2015.

[8] A. Wibisono, M. J. Wainwright, M. Jordan, and J. C. Duchi, "Finite sample convergence rates of zero-order stochastic optimization methods," in *Advances in Neural Information Processing Systems*, 2012.

[9] H. J. Kushner, "A New Method of Locating the Maximum Point of an Arbitrary Multipeak Curve in the Presence of Noise," *Journal of Basic Engineering*, vol. 86, no. 1, pp. 97–106, 1964.

[10] J. Močkus, "On bayesian methods for seeking the extremum," in *Optimization Techniques IFIP Technical Conference Novosibirsk, July 1–7, 1974*, pp. 400–404, Springer Berlin Heidelberg, 1975.

[11] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[12] P. I. Frazier, "A tutorial on bayesian optimization," 2018.

[13] Y. Ha and S. Shashaani, "Iteration complexity and finite-time efficiency of adaptive sampling trust-region methods for stochastic derivative-free optimization," *IISE Transactions*, vol. 0, no. 0, pp. 1–15, 2024.

[14] N. Srinivas, A. Krause, S. Kakade, and M. Seeger, "Gaussian process optimization in the bandit setting: no regret and experimental design," in *Proceedings of the 27th International Conference on International Conference on Machine Learning*, p. 1015–1022, 2010.

[15] W. R. Thompson, "On the likelihood that one unknown probability exceeds another in view of the evidence of two samples," *Biometrika*, vol. 25, no. 3/4, pp. 285–294, 1933.

[16] D. Russo, B. V. Roy, A. Kazerouni, I. Osband, and Z. Wen, "A tutorial on thompson sampling," 2020.

[17] O. Chapelle and L. Li, "An empirical evaluation of thompson sampling," in *Advances in Neural Information Processing Systems*, vol. 24, Curran Associates, Inc., 2011.

[18] S. Vakili, N. Bouziani, S. Jalali, A. Bernacchia, and D.-s. Shiu, "Optimal order simple regret for gaussian process bandits," in *Proceedings of the 35th International Conference on Neural Information Processing Systems*, 2024.

[19] T. Lai and H. Robbins, "Asymptotically efficient adaptive allocation rules," *Advances in Applied Mathematics*, vol. 6, no. 1, pp. 4–22, 1985.

[20] S. Agrawal and N. Goyal, "Thompson sampling for contextual bandits with linear payoffs," in *Proceedings of the 30th International Conference on Machine Learning*, vol. 28, pp. 127–135, PMLR, 2013.

[21] H. Ma, T. Zhang, Y. Wu, F. P. Calmon, and N. Li, "Gaussian max-value entropy search for multi-agent bayesian optimization," 2023.

[22] K. Kandasamy, A. Krishnamurthy, J. Schneider, and B. Poczos, "Parallelised bayesian optimisation via thompson sampling," in *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, vol. 84, pp. 133–142, PMLR, 2018.

[23] J. Wu and P. Frazier, "The parallel knowledge gradient method for batch bayesian optimization," in *Advances in Neural Information Processing Systems*, vol. 29, Curran Associates, Inc., 2016.

[24] Z. Ren and N. Li, "Minimizing the thompson sampling regret-to-sigma ratio (ts-rsr): a provably efficient algorithm for batch bayesian optimization," *arXiv preprint arXiv:2403.04764*, 2024.

[25] J. M. Hernández-Lobato, J. Requeima, E. O. Pyzer-Knapp, and A. Aspuru-Guzik, "Parallel and distributed thompson sampling for large-scale accelerated exploration of chemical space," in *Proceedings of the 34th International Conference on Machine Learning*, vol. 70, PMLR, 2017.

[26] T. Zhang, V. Qin, Y. Tang, and N. Li, "Distributed information-based source seeking," *IEEE Transactions on Robotics*, vol. 39, no. 6, pp. 4749–4767, 2023.

[27] K. Nakamura, M. Santos, and N. E. Leonard, "Decentralized learning with limited communications for multi-robot coverage of unknown spatial fields," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 9980–9986, 2022.

[28] J. Garcia-Barcos and R. Martinez-Cantin, "Fully distributed bayesian optimization with stochastic policies," in *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*, International Joint Conferences on Artificial Intelligence Organization, 2019.

[29] M. Jamil, X.-S. Yang, and H.-J. Zepernick, "8 - test functions for global optimization: A comprehensive survey," in *Swarm Intelligence and Bio-Inspired Computation*, pp. 193–222, Elsevier, 2013.

[30] D. Russo and B. V. Roy, "An information-theoretic analysis of thompson sampling," *Journal of Machine Learning Research*, vol. 17, no. 68, pp. 1–30, 2016.

[31] S. Zerefa, Z. Ren, H. Ma, and N. Li, "Distributed thompson sampling under constrained communication," *arXiv preprint arXiv:2410.15543*, 2024.

[32] M. Balandat, B. Karrer, D. Jiang, S. Daulton, B. Letham, A. G. Wilson, and E. Bakshy, "Botorch: A framework for efficient monte-carlo bayesian optimization," in *Advances in Neural Information Processing Systems*, vol. 33, pp. 21524–21538, Curran Associates, Inc., 2020.

[33] P. Erdős and A. Rényi, "On random graphs i," *Publ. math. debrecen*, vol. 6, no. 290-297, p. 18, 1959.

APPENDIX I
ADDITIONAL ANALYSIS

Lemma I.1. *([22] [30]) At step j, for all $x \in \mathcal{A}$, $\mathbb{E}[\mathbb{I}\{f(x) > U_j(x)\} \cdot (f(x) - U_j(x))] \leq \frac{1}{\sqrt{2\pi}} e^{-\beta_j/2}$.*

Proof. Since f is a GP, we know $f(x)|D_j \sim \mathcal{N}(\mu_j(x), \sigma_j^2(x))$. Recall that $U_j(\cdot) = \mu_j(\cdot) + \beta_j^{1/2}\sigma_j(\cdot)$. Thus, we know that $f(x) - U_j(x)|D_j \sim \mathcal{N}(-\beta_j^{1/2}\sigma_j(x), \sigma_j^2(x))$. For a normal distribution $Z \sim \mathcal{N}(\mu, \sigma^2)$, with $\mu \leq 0$, we have that $\mathbb{E}[Z\mathbb{I}(Z > 0)] \leq \frac{\sigma}{\sqrt{2\pi}} e^{-\mu^2/(2\sigma^2)}$. Thus, we can apply this fact to our setting, yielding:

$$\begin{aligned}
\mathbb{E}[\mathbb{I}\{f(x) > U_j(x)\} \cdot (f(x) - U_j(x))] &\leq \frac{\sigma_j(x)}{\sqrt{2\pi}} e^{-\beta_j/2} && \text{(Aforementioned property of } \mathcal{N}) \\
&\leq \frac{1}{\sqrt{2\pi}} e^{-\beta_j/2} && (\sigma_j(x) \leq \kappa(x,x) \leq 1)
\end{aligned}$$

□

Lemma I.2. *[14] The information gain for selected points can be expressed in terms of the predictive variances. if $f_{[n]} = (f(x_n)) \in \mathbb{R}^n$:*

$$I(y_{[n]}; f_{[n]}) = \frac{1}{2} \sum_{j=1}^{n} \log(1 + \sigma_\epsilon^{-2} \sigma_{j-1}^2(x_n))$$

Proof.

$$\begin{aligned}
I(y_{[n]}; f_{[n]}) &= H(y_{[n]}) + H(y_{[n]}|f_{[n]}) && \text{(Def of information)} \\
&= H(y_{[n]}) - \frac{1}{2}\log|2\pi e\sigma_\epsilon^2 \mathbf{I}| && \text{(Gaussian entropy: } H(N(\mu,\Sigma)) = \frac{1}{2}\log|2\pi e\Sigma|)
\end{aligned}$$

Now let's develop an expression for $H(y_{[n]})$:

$$H(y_{[n]}) = H(y_{[n-1]}) + H(y_n|y_{[n-1]}) \qquad \text{(Entropy chain rule: } H(A,B) = H(A) + H(B|A))$$

$$= H(y_{[n-1]}) + \frac{1}{2}\log(2\pi e(\sigma_\epsilon^2 + \sigma_{n-1}^2(x_n))/2 \quad (y_n|y_{[n-1]} = f(x_n) + \epsilon_n, y_n|y_{[n-1]} \sim N(\mu_{n-1}(x_n), \sigma_{n-1}^2(x_n) + \sigma_\epsilon^2)$$

We now have developed a recursive relation for $H(y_{[n]})$; we can inductively show that $H(y_{[n]}) = \sum_{j=1}^{n} \frac{1}{2}\log(2\pi e(\sigma_\epsilon^2 + \sigma_{j-1}^2(x_n))/2$. Utilizing this expression, we can return to our mutual information expression:

$$\begin{aligned}
I(y_{[n]}; f_{[n]}) &= H(y_{[n]}) - \frac{1}{2}\log|2\pi e\sigma_\epsilon^2\mathbf{I}| && \text{(Previously shown)} \\
&= \sum_{j=1}^{n} \frac{1}{2}\log(2\pi e(\sigma_\epsilon^2 + \sigma_{j-1}^2(x_n))/2 - \frac{1}{2}\log|2\pi e\sigma_\epsilon^2\mathbf{I}| && \text{(Substituting developed } H(y_{[n]} \text{ expression)} \\
&= \frac{1}{2}\sum_{j=1}^{n} \log(1 + \sigma_{j-1}^2(x_n)\sigma_\epsilon^{-2}) && \text{(Simplifying log subtraction)}
\end{aligned}$$

□

Lemma I.3. *[22] Let $f \sim GP(0, \kappa), f : \mathcal{X} \to \mathbb{R}$ and each time we query $x \in \mathbb{X}$ we observe $y = f(x) + \varepsilon$, where $\varepsilon \sim \mathcal{N}(0, \sigma_n^2)$. Let $\{x_1, \ldots, x_t\}$ be an arbitrary set of t evaluations to f where $x_j \in \mathcal{X}$ for all j. Let σ_j^2 denote the posterior variance conditioned on the first j of these queries, $\{x_1, \ldots, x_j\}$. Then, $\sum_{j=1}^{n} \sigma_{j-1}^2(x_j) \leq \frac{2\Psi_n}{\log(1+\sigma_\epsilon^{-2})}$.*

Proof.

$$\begin{aligned}
\Psi_n &\geq I(y_{[n]}; f_{[n]}) && \text{(Definition of } \Psi_n) \\
&\geq \frac{1}{2}\sum_{j=1}^{n} \log(1 + \sigma_\epsilon^{-2}\sigma_{j-1}^2(x_j)) && \text{(Lemma I.2)}
\end{aligned}$$

Note that the function $\frac{x}{1+\log(x)}$ increases with x. Also note that $0 \leq \sigma_\epsilon^2, \sigma_j^2(x_j) \leq 1$ by assumption, and thus $\sigma_\epsilon^{-2}\sigma_j^2(x_j) \leq \sigma_\epsilon^{-2}$. Therefore,

$$\frac{\sigma_\epsilon^{-2}\sigma_{j-1}^2(x_j)}{\log(1 + \sigma_\epsilon^{-2}\sigma_{j-1}^2(x_j))} \leq \frac{\sigma_\epsilon^{-2}}{\log(1+\sigma_\epsilon^{-2})} \implies \sigma_{j-1}^2(x_j)\log(1+\sigma_\epsilon^{-2}) \leq \log(1 + \sigma_\epsilon^{-2}\sigma_{j-1}^2(x_j)).$$

Now we can revisit our earlier Ψ_n expression:

$$\begin{aligned}\Psi_n &\geq \frac{1}{2}\sum_{j=1}^{n}\log(1+\sigma_\epsilon^{-2}\sigma_{j-1}^2(x_j)) && \text{(Lemma I.2)}\\ &\geq \frac{1}{2}\log(1+\sigma_\epsilon^{-2})\sum_{j=1}^{n}\sigma_{j-1}^2(x_j) && \text{(Previously shown)}\end{aligned}$$

With algebraic manipulation, the last expression is equivalent to $\sum_{j=1}^{n}\sigma_{j-1}^2(x_j) \leq \frac{2\Psi_n}{\log(1+\sigma_\epsilon^{-2})}$, our desired statement. □

Lemma I.4. *Let $f \sim GP(0,\kappa)$, and let A, B be finite subsets of $\mathcal{X}$. Let $y_A \in \mathbb{R}^{|A|}$ and $y_B \in \mathbb{R}^{|B|}$ denote the observations when we evaluate f at A and B. Let $\sigma_A, \sigma_{A\cup B} : \mathcal{X} \to \mathbb{R}$ denote the posterior standard deviation of the GP when conditioned on A and $A \cup B$, respectively. Then*

$$\forall x \in \mathcal{X}, \qquad \frac{\sigma_A(x)}{\sigma_{A\cup B}(x)} = exp(I(f; y_B|y_A))$$

Proof.

$$\begin{aligned}I(f; y_B|y_A) &= H(f|y_A) - H(f|y_{A\cup B}) && \text{(Mutual information with conditional entropy)}\\ &= \frac{1}{2}\log|2\pi e\sigma_A^2| - \frac{1}{2}\log|2\pi e\sigma_{A\cup B}^2| && \text{(Entropy for GP)}\\ &= \log\left(\frac{\sigma_A}{\sigma_{A\cup B}}\right)\end{aligned}$$

Thus, $\frac{\sigma_A}{\sigma_{A\cup B}} = exp(I(f; y_B|y_A))$ □

APPENDIX II
BOUNDS FOR THE INFORMATION GAIN QUANTITY Ψ_τ FOR DIFFERENT KERNELS

We note that following a known result in [14], Ψ_τ in fact satisfies sublinear growth for three well-known classes of kernels, namely the linear, exponential and Matern kernels.

Lemma II.1 (cf. Theorem 5 in [14])**.** *For any $\tau > 0$, the maximal information gain Ψ_τ can be bounded as follows for the following kernels.*

1) *(Linear kernel): If $k(x, x') = x^\top x'$, then*
$$\Psi_\tau = O(d\log(\tau)).$$
2) *(Squared exponential kernel): If $k(x, x') = \exp(-\|x - x'\|^2/2)$, then*
$$\Psi_\tau = O((\log(\tau))^{d+1}).$$
3) *(Matern kernel with $\nu > 1$): If $k(x, x') = \frac{1}{\Gamma(\nu)2^{\nu-1}}\left(\frac{\sqrt{2\nu}}{d}\|x - x'\|\right)^v K_v\left(\frac{\sqrt{2v}}{d}\|x - x'\|\right)$, where $K_v(\cdot)$ is a modified Bessel function, and $\Gamma(\cdot)$ denotes the gamma function, then*
$$\Psi_\tau = O((\tau)^{\frac{d(d+1)}{2\nu+d(d+1)}}\log(\tau))$$

APPENDIX III
OBJECTIVE FUNCTIONS

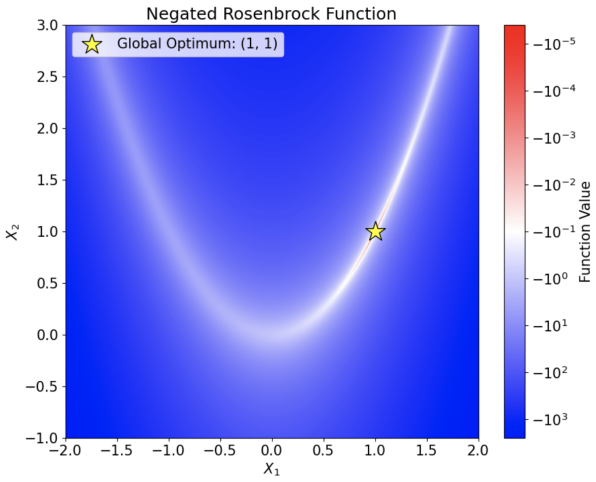
The test objective functions used in our simulations were the Rosenbrock and Ackley functions. To align with the maximization objectives of the distributed Thompson sampling algorithm, we negated the canonical definitions of the Rosenbrock and Ackley functions so that the global optima would be a maximum. The equations of the modified objective functions are as follows:

Negated Rosenbrock:

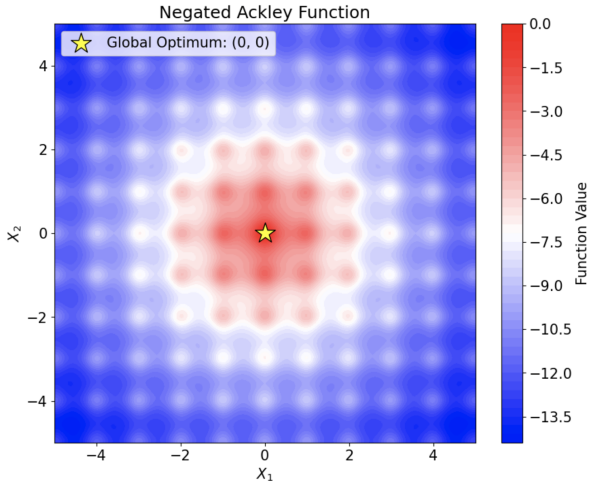
$$f(x_1, x_2) = -(1 - x_1)^2 - 100(x_2 - x_1^2)^2 \tag{18}$$

Negated Ackley:

$$f(x_1, x_2) = 20\exp\left(-0.2\sqrt{\frac{x_1^2 + x_2^2}{2}}\right) + \exp\left(\frac{1}{2}(\cos(2\pi x_1) + \cos(2\pi x_2))\right) - 20 - \exp(1) \tag{19}$$



(a) Plot of Negated Rosenbrock Function.



(b) Plot of Negated Ackley Function.

Fig. 2: Plots of test objective functions. The location of the global optimum is denoted by a yellow star.