
Asynchronous ϵ -Greedy Bayesian Optimisation

George De Ath¹

Richard M. Everson¹

Jonathan E. Fieldsend¹

¹Department of Computer Science, University of Exeter, Exeter, United Kingdom

Abstract

Batch Bayesian optimisation (BO) is a successful technique for the optimisation of expensive black-box functions. Asynchronous BO can reduce wallclock time by starting a new evaluation as soon as another finishes, thus maximising resource utilisation. To maximise resource allocation, we develop a novel asynchronous BO method, AEGiS (Asynchronous ϵ -Greedy Global Search) that combines greedy search, exploiting the surrogate’s mean prediction, with Thompson sampling and random selection from the approximate Pareto set describing the trade-off between exploitation (surrogate mean prediction) and exploration (surrogate posterior variance). We demonstrate empirically the efficacy of AEGiS on synthetic benchmark problems, meta-surrogate hyperparameter tuning problems and real-world problems, showing that AEGiS generally outperforms existing methods for asynchronous BO. When a single worker is available performance is no worse than BO using expected improvement.

1 INTRODUCTION

Bayesian Optimisation (BO) is a popular sequential approach for optimising time-consuming or costly black-box functions that have no closed-form expression or derivative information [Brochu et al., 2010, Snoek et al., 2012]. BO comprises two main steps which are repeated until convergence or the budget is exhausted. Firstly, a surrogate model is constructed using previous evaluations; this is typically a Gaussian process (GP), due to its strength in uncertainty quantification and function approximation [Rasmussen and Williams, 2006, Shahriari et al., 2016]. Secondly, an acquisition function is optimised to select the next location to expensively evaluate. Such functions attempt to balance the

exploitation of locations with a good predicted value with the exploration of locations with high uncertainty.

In real-world problems it is often possible to use hardware capabilities to run multiple evaluations in parallel. For example, the number of hyperparameter configurations that can be evaluated in parallel when training neural networks [Chen et al., 2018] is only limited by the computational resources available. The desire to speed up optimisation wallclock time through parallel evaluation led to the development of parallel BO in which q locations are jointly selected to be evaluated synchronously. It has been successfully applied in many areas such as drug discovery [Hernández-Lobato et al., 2017], heat treatment, [Gupta et al., 2018], and hyperparameter configuration [Kandasamy et al., 2018].

When the runtime for evaluations varies, such as when optimising the number of units in the layers of a neural network, or indeed the number of layers themselves, synchronous parallel approaches will result in resource underutilisation because all evaluations must finish before the next q locations can be suggested. To counteract this, asynchronous BO approaches have been proposed that allow for the evaluation of functions asynchronously; i.e., as soon as an evaluation has completed, another can be submitted. In order to take into account the uncertainty in the surrogate model at and around locations which are still being evaluated, a number of methods have been proposed to penalise either the surrogate model itself or the acquisition function in order to limit the selection of locations near places that are currently under evaluation [Ginsbourger et al., 2008, González et al., 2016, Alvi et al., 2019].

Kandasamy et al. [2018] presented an alternative approach based on Thompson sampling (TS) that draws surrogate function realisations from the GP posterior and optimises these to suggest new locations to evaluate without the need to penalise the model or acquisition function. However, TS relies on there being enough stochasticity to create sufficiently diverse function draws with meaningfully different optima. Therefore, TS tends to be insufficiently exploratory

in low dimensions (as draws are very similar to the mean prediction) and over-exploratory in high dimensions where the posterior uncertainty is large across much of the domain.

ϵ -greedy methods have been shown to be remarkably successful in sequential and synchronous parallel optimisation [De Ath et al., 2020, De Ath et al., 2021]. Most of the time these exploit the mean prediction of the surrogate model by expensively evaluating the location with the best posterior mean, ignoring the uncertainty in the surrogate’s prediction, but making deliberately exploratory samples with probability ϵ . In this work we combine a greedy exploitation of the mean prediction with Thompson sampling and random selection from the approximate Pareto set that maximally trades off exploration and exploitation. Specifically, we introduce an ϵ -greedy scheme that, with probability ϵ_T performs Thompson sampling or with probability ϵ_P random selection from the approximate Pareto set, and exploits the surrogate model’s posterior mean prediction the rest of the time, i.e. with probability $1 - (\epsilon_T + \epsilon_P)$. For convenience we write $\epsilon = \epsilon_T + \epsilon_P$. One of the reasons for the success of ϵ -greedy methods is that model inaccuracy introduces some accidental exploration even when exploiting the mean prediction. This effect becomes dominant as the dimension increases, and we therefore investigate how ϵ can be reduced with increasing dimension.

- We present AEGiS (Asynchronous ϵ -Greedy Global Search), a novel asynchronous BO algorithm that combines greedy exploitation of the posterior mean prediction with Thompson sampling and random selection from the approximate Pareto set between exploration and exploitation.
- We present results on fifteen well-known benchmark functions, three meta-surrogate hyperparameter tuning problems and two robot pushing problems. Empirically, we demonstrate that AEGiS outperforms existing methods for asynchronous BO on synthetic and hyperparameter optimisation problems.
- We investigate how the degree of deliberate exploration ϵ should be controlled with increasing dimension, d , showing empirically that $\epsilon = \min(2/\sqrt{d}, 1)$ yields superior performance to increased or decreased rates of exploration with respect to problem dimensionality.
- We present an ablation study that verifies the importance of all three components, namely the Pareto set selection, Thompson sampling, and exploitation, showing that AEGiS achieves performance that is greater than the sum of its individual components.
- We show empirically that selection from the exploration-exploitation Pareto set is superior to selection from the entire problem domain.

We begin in Section 2 by reviewing sequential, synchronous and asynchronous BO, and Thompson sampling, before

introducing AEGiS in Section 3. Section 4 details extensive empirical evaluation of AEGiS against other popular methods on well-known test problems, and also includes an ablation study. We finish with concluding remarks in Section 5.

2 BAYESIAN OPTIMISATION

Our goal is to minimise a black-box function $f : \mathcal{X} \mapsto \mathbb{R}$, defined on a compact domain $\mathcal{X} \subset \mathbb{R}^d$. The function itself is unknown, but we have access to the result of its evaluation $f(\mathbf{x})$ at any location $\mathbf{x} \in \mathcal{X}$. We are particularly interested in cases where the evaluations are expensive, either in terms of time or money or both. We therefore seek to minimise f in either as few evaluations as possible to incur as little cost as possible or for a fixed budget.

2.1 SEQUENTIAL BAYESIAN OPTIMISATION

Bayesian Optimisation (BO) is a global search strategy that sequentially samples the problem domain \mathcal{X} at locations that are likely to contain the global optimum, taking into account both the predictions of a probabilistic surrogate model and its corresponding uncertainty [Jones et al., 1998]. It starts by generating M initial sample locations $\{\mathbf{x}_i\}_{i=1}^M$ with a space filling algorithm, such as Latin hypercube sampling [McKay et al., 1979], and expensively evaluates them with the function $f_i = f(\mathbf{x}_i)$. This set of observations $\mathcal{D}_M = \{(\mathbf{x}_n, f_n \triangleq f(\mathbf{x}_n))\}_{n=1}^M$ is used for initial training of the surrogate model. Following training, and at each iteration of BO, the next location to be expensively evaluated is chosen as the location that maximises an acquisition function $\alpha(\mathbf{x})$. Thus $f(\cdot)$ is next evaluated at $\mathbf{x}' = \operatorname{argmax}_{\mathbf{x} \in \mathcal{X}} \alpha(\mathbf{x})$. The dataset is augmented with \mathbf{x}' and $f(\mathbf{x}')$ and the process is repeated until the budget is exhausted. The value of the global minimum is then estimated to be the best function evaluation seen during optimisation, i.e. $f^* = \min_i \{f_i\}$.

Gaussian Processes Gaussian processes (GPs) are commonly used as the surrogate model for f . A GP defines a prior distribution over functions, such that any finite number of function values have a joint Gaussian distribution [Rasmussen and Williams, 2006] with mean $m(\mathbf{x})$ and a covariance $\kappa(\mathbf{x}, \mathbf{x}' | \boldsymbol{\theta})$ with hyperparameters $\boldsymbol{\theta}$. Henceforth, w.l.o.g. we use a zero-mean prior $m(\mathbf{x}) = 0 \forall \mathbf{x} \in \mathcal{X}$. Conditioning the prior distribution on data consisting of f evaluated at t sampled locations \mathcal{D}_t , the posterior distribution of f is also a GP with posterior mean and variance

$$\mu(\mathbf{x} | \mathcal{D}, \boldsymbol{\theta}) = m(\mathbf{x}) + \boldsymbol{\kappa}(\mathbf{x}, X)K^{-1}\mathbf{f} \quad (1)$$

$$\sigma^2(\mathbf{x} | \mathcal{D}, \boldsymbol{\theta}) = \kappa(\mathbf{x}, \mathbf{x}) - \boldsymbol{\kappa}(\mathbf{x}, X)^\top K^{-1}\boldsymbol{\kappa}(X, \mathbf{x}). \quad (2)$$

Here, $X \in \mathbb{R}^{t \times d}$ and $\mathbf{f} = (f_1, f_2, \dots, f_t)^\top$ are the matrix of design locations and corresponding vector of function evaluations respectively. The kernel matrix $K \in$

$\mathbb{R}^{t \times t}$ is $K_{ij} = \kappa(\mathbf{x}_i, \mathbf{x}_j | \boldsymbol{\theta})$ and $\kappa(\mathbf{x}, X) \in \mathbb{R}^t$ is given by $[\kappa(\mathbf{x}, X)]_i = \kappa(\mathbf{x}, \mathbf{x}_i | \boldsymbol{\theta})$. Kernel hyperparameters $\boldsymbol{\theta}$ are learnt by maximising the log marginal likelihood [Rasmussen and Williams, 2006].

Acquisition Functions The choice of where to evaluate next in BO is determined by an acquisition function $\alpha : \mathbb{R}^d \mapsto \mathbb{R}$. Perhaps the most commonly used acquisition function is Expected Improvement (EI) [Jones et al., 1998], which is defined as the expected positive improvement over the best-seen function value f^* thus far:

$$\alpha_{\text{EI}}(\mathbf{x}) = \mathbb{E}_{p(f(\mathbf{x}) | \mathcal{D})} [\max(f^* - f(\mathbf{x}), 0)]. \quad (3)$$

However, many other acquisition functions have also been proposed, including probability of improvement [Kushner, 1964], optimistic strategies such as UCB [Srinivas et al., 2010], ϵ -greedy strategies [Bull, 2011, De Ath et al., 2021], and information-theoretic approaches [Scott et al., 2011, Hennig and Schuler, 2012, Henrandez-Lobato et al., 2014, Wang and Jegelka, 2017, Ru et al., 2018].

2.2 PARALLEL BAYESIAN OPTIMISATION

In parallel BO, the algorithm has access to q workers that can evaluate f in parallel and the focus is on minimising the optimisation time, acknowledging that the total computational cost will be greater than for strictly sequential evaluations. The goal is to select a batch of q promising locations to be expensively evaluated in parallel. In the synchronous setting, the BO algorithm sends out a batch of q locations to be evaluated in parallel, and waits for all q evaluations to be completed before submitting the next batch. However, in the asynchronous setting, as soon as a worker has finished evaluating f a new location is submitted to it for evaluation. Therefore, if f takes a variable time to evaluate, the asynchronous version of BO will result in more evaluations of f for the same wallclock time.

One of the earliest synchronous approaches was the qEI method of Ginsbourger et al. [2008], a generalisation of the sequential EI acquisition function (3) that jointly proposes a batch of q locations to evaluate. Similarly, the parallel predictive entropy search [Shah and Ghahramani, 2015] and the parallel knowledge gradient [Wu and Frazier, 2016] also jointly propose q locations and are based on their sequential counterparts. However, these methods require the optimisation of a $d \times q$ problem that becomes prohibitively expensive to solve as q increases [Daxberger and Low, 2017].

Consequently, the prevailing strategy has become to sequentially select the q batch locations. Penalisation-based methods penalise the regions of either the surrogate model [Azimi et al., 2010, Desautels et al., 2014] or the acquisition function that correspond to locations that are currently under evaluation. The Kriging Believer method of Ginsbourger et al. [2010], for example, hallucinates the results of pending

evaluations by using the surrogate model’s mean prediction as its observed value. The local penalisation methods of Gonzalez et al. [2016] and Alvi et al. [2019] directly penalise an acquisition function. De Ath et al. [2020] show that an ϵ -greedy method that usually exploits the most promising region, with occasional exploratory moves, is effective.

Asynchronous BO has received less attention. Janusevskis et al. [2012] proposed an asynchronous version of the qEI method that uses the qEI criterion assuming that the p locations currently under evaluation are fixed, and optimises the position of the remaining $q - p$ locations. Kandasamy et al. [2018] suggest using Thompson sampling to select locations to evaluate. Thompson sampling (TS) [Thompson, 1933] is a randomised strategy for sequential decision-making under uncertainty [Kandasamy et al., 2018]. At each iteration, TS selects the next location to evaluate \mathbf{x}' according to the posterior probability that it is the optimum. If $p_*(\mathbf{x} | \mathcal{D})$ denotes the probability that \mathbf{x} minimises the surrogate model posterior, then

$$p_*(\mathbf{x} | \mathcal{D}) = \int p_*(\mathbf{x} | g) p(g | \mathcal{D}) dg \quad (4)$$

$$= \int \delta(\mathbf{x} - \underset{\mathbf{x} \in \mathcal{X}}{\operatorname{argmin}} g(\mathbf{x})) p(g | \mathcal{D}) dg, \quad (5)$$

where δ is the Dirac delta distribution. Thus, all the probability mass of $p_*(\mathbf{x} | \mathcal{D})$ is at the minimiser of g . In the context of BO, this corresponds to sampling a function $g(\cdot)$ from the surrogate model’s posterior $p(\cdot | \mathbf{x}, \mathcal{D})$ and selecting $\mathbf{x}' = \underset{\mathbf{x} \in \mathcal{X}}{\operatorname{argmin}} g(\mathbf{x})$ as the next location to evaluate. This can then be trivially applied in the sequential, synchronous and asynchronous settings by sampling and minimising as many function realisations as required. Thus, the BO loop using TS involves building a surrogate model with previously evaluated locations (ignoring those that may be currently under evaluation), drawing as many function realisations as there are free workers, finding the minimiser of each realisation and evaluating f at those locations.

Kandasamy et al. [2018] analyse Thompson sampling in terms of maximum information gain. They show that both synchronous and asynchronous TS algorithms making n function evaluations are almost as good as if the n evaluations were made sequentially. Furthermore, they extend the notion of simple regret to incorporate the (random) number of evaluations made in a given time. Under this metric they show that asynchronous TS outperforms synchronous and sequential algorithms using a variety of evaluation-time models.

Traditionally, g has been approximately optimised by picking the minimum of a randomly sampled set of discrete points [Kandasamy et al., 2018]. Recently, Wilson et al. [2020] have identified an elegant decoupled sampling method, which we use here, that decomposes the GP posterior into a sum of a weight-space prior term that is approximated by random Fourier features, and a pathwise update

in function space. Unlike other fast approximations, such as [Rahimi and Recht, 2008], this method does not suffer from variance starvation [Calandriello et al., 2019]. It scales linearly with the number of locations queried during optimisation of g , and is pathwise differentiable, meaning that gradient-based optimisation can be employed.

Recently, De Palma et al. [2019] proposed to draw realisations of acquisition functions instead, by drawing a new set of surrogate model hyperparameters from their posterior distribution for each acquisition function. An advantage of these TS-based methods is that each draw from either the model or hyperparameter posterior defines a new distinct acquisition function. A drawback is that they rely on the uncertainty in the surrogate model to give sufficient diversity in position to the locations selected, unlike acquisition functions which implicitly or explicitly balance exploration and exploitation.

3 AEGIS

It is well recognised that the fundamental problem in global optimisation is to balance exploitation of promising locations found thus far with exploration of new areas which have not been explored. While extremely appealing in its simplicity, Thompson sampling (TS) tends to over-exploit in low dimensions (because after a number of evaluations the posterior variance is reduced and most sampled realisations are similar to the posterior mean), but under-exploit in high dimensions (because there are insufficient samples to drive the posterior variance away from the prior, resulting in effectively random search). The efficacy of ϵ -greedy methods in sequential and synchronous BO over a range of dimensions [De Ath et al., 2020, De Ath et al., 2021] motivates an ϵ -greedy strategy to switch between greedy exploitation of the surrogate model, exploration, and TS. Specifically, at each step we optimise a TS function draw with probability ϵ_T or perform exploration each with probability ϵ_P , and exploit the surrogate model’s posterior mean prediction $\mu(\mathbf{x})$ of the rest of the time, i.e. with probability $1 - (\epsilon_T + \epsilon_P)$.

The AEGiS method is outlined in Algorithm 1. With probability $1 - (\epsilon_T + \epsilon_P)$, the mean prediction from the surrogate model is minimised to find the next location \mathbf{x}' to evaluate (line 6). Alternatively, with probability ϵ_T , a TS step taken: the next location is determined by sampling a realisation $g(\cdot)$ from the surrogate posterior and finding its minimiser (lines 8 and 9). Otherwise (lines 11 and 12), with probability ϵ_P , the location to evaluate is chosen from the approximate Pareto set of all locations which maximally trade off between exploitation (minimising $\mu(\mathbf{x})$) and exploration (maximising $\sigma^2(\mathbf{x})$). A full discussion is given by De Ath et al. [2021]. Briefly, a location \mathbf{x}_1 dominates \mathbf{x}_2 (written $\mathbf{x}_1 \succ \mathbf{x}_2$) iff $\mu(\mathbf{x}_1) \leq \mu(\mathbf{x}_2)$ and $\sigma(\mathbf{x}_1) \geq \sigma(\mathbf{x}_2)$, and at least one of the inequalities is strict. If both $\mathbf{x}_1 \not\succeq \mathbf{x}_2$ and $\mathbf{x}_2 \not\succeq \mathbf{x}_1$ then \mathbf{x}_1 and \mathbf{x}_2 are said

Algorithm 1 AEGiS

Inputs: \mathcal{D}_t : Training data; T : Evaluation budget

- 1: $M = |\mathcal{D}_t|$
- 2: **for** $t = M + 1, \dots, T$ **do**
- 3: Condition model posterior on \mathcal{D}_t ; see (1) and (2)
- 4: $r \sim \mathcal{U}(0, 1)$
- 5: **if** $r < 1 - (\epsilon_T + \epsilon_P)$ **then** \triangleright Exploit surrogate model
- 6: $\mathbf{x}' \leftarrow \underset{\mathbf{x} \in \mathcal{X}}{\operatorname{argmin}} \mu(\mathbf{x})$
- 7: **else if** $r < 1 - \epsilon_P$ **then** \triangleright Thompson sample
- 8: $g \sim p(f | \mathcal{D})$
- 9: $\mathbf{x}' \leftarrow \underset{\mathbf{x} \in \mathcal{X}}{\operatorname{argmin}} g(\mathbf{x})$
- 10: **else** \triangleright Approximate Pareto set selection
- 11: $\tilde{\mathcal{P}} \leftarrow \operatorname{M0Optimise}_{\mathbf{x} \in \mathcal{X}}(\mu(\mathbf{x}), \sigma^2(\mathbf{x}))$
- 12: $\mathbf{x}' \leftarrow \operatorname{randomChoice}(\tilde{\mathcal{P}})$
- 13: Submit new job to evaluate \mathbf{x}'
- 14: **while** all workers in use **do**
- 15: Wait for a worker to finish
- 16: $\mathcal{D}_t \leftarrow \mathcal{D}_{t-1} \cup \{(\mathbf{x}', f(\mathbf{x}'))\}$ \triangleright Augment data

to be mutually non-dominating. The maximal set of mutually non-dominating solutions is known as the Pareto set $\mathcal{P} = \{\mathbf{x} \in \mathcal{X} \mid \mathbf{x}' \not\succeq \mathbf{x} \forall \mathbf{x}' \in \mathcal{X}\}$ and comprises the optimal trade-off between exploratory and exploitative locations. Since it is infeasible to enumerate \mathcal{X} , we sample a location from the approximate Pareto set $\tilde{\mathcal{P}} \approx \mathcal{P}$ which is found using a standard multi-objective optimiser (lines 11 and 12).

Selecting a random location from $\tilde{\mathcal{P}}$ ensures that no other location could have been selected that has either more uncertainty given the same predicted value or a lower predicted value given the same amount of uncertainty. Therefore, when considering the two objectives of both reducing the overall model uncertainty and finding the location with the minimum value, these locations are optimal with respect to a specific exploitation-exploration trade-off. Indeed, the maximisers of both EI and UCB are members of \mathcal{P} [De Ath et al., 2021].

Rather than selecting from $\tilde{\mathcal{P}}$, a more exploratory procedure is to select from the entire feasible space \mathcal{X} . That is, lines 11 and 12 are replaced with $\mathbf{x}' \leftarrow \operatorname{randomChoice}(\mathcal{X})$. We denote this variant by AEGiS-RS. However, as shown below, it is less effective than selection from $\tilde{\mathcal{P}}$.

At the beginning of an optimisation run, the initial q locations to be evaluated are chosen by first selecting the most exploitative location (line 6). The remaining $q - 1$ locations are each chosen by performing either Thompson sampling or selection from the approximate Pareto set $\tilde{\mathcal{P}}$ with probabilities $\epsilon_T / (\epsilon_T + \epsilon_P)$ and $\epsilon_P / (\epsilon_T + \epsilon_P)$ respectively. This ensures that the most exploitative location is only chosen once during initialisation. The remaining evaluated locations in an optimisation run are chosen according to Algorithm 1.

As discussed above, inaccuracies in the surrogate model provide an element of exploration even when disregarding

the uncertainty in the posterior $p(f | \mathcal{D}_t)$. On the other hand, TS tends to be over-exploitative in low dimensions ($d \lesssim 4$) and too exploratory in high dimensions. The performance of sequential ϵ -greedy algorithms is relatively insensitive to the precise choice of ϵ [De Ath et al., 2021]. As we show in section 4.4, the performance of AEGiS is relatively insensitive to the ratio of $\epsilon_T : \epsilon_P$. For simplicity, we therefore initially set $\epsilon_T = \epsilon_P = \epsilon/2$ and $\epsilon = \min(2/\sqrt{d}, 1)$. This allows for more exploitation to occur as the dimensionality of the problem increases. As shown empirically in section 4.4, this rate of decay in the amount of exploration carried out with respect to the problem dimensionality is more effective than faster or slower rates.

The dominant contributor to computational cost of finding the approximate Pareto set, carried out by NSGA-II [Deb et al., 2002], is the non-dominated sort of the population of solutions and their offspring. With two objectives, as is the case in AEGiS, this has complexity $\mathcal{O}(GN \log N)$ [Jensen, 2003], where N is the total population size and G is the number of generations for which the evolutionary algorithm runs. Empirically we find that this cost is similar to that of TS, whose computational cost is dominated by the gradient descent method used to optimise the sample path g ; e.g. $\mathcal{O}(Ld^2)$ for BFGS [Byrd et al., 1995], where L is the maximum number of steps taken in the search. However, we note that both of these costs are negligible compared to fitting the GP.

4 EXPERIMENTAL EVALUATION

We compare AEGiS with four well-known asynchronous BO algorithms: two acquisition function-based penalisation methods, Local Penalisation (LP) [González et al., 2016] and the more recent PLAyBOOK [Alvi et al., 2019], as well as the Kriging Believer (KB) [Ginsbourger et al., 2010], all of which are based on the expected improvement acquisition function (3). We also include the standard Thompson sampling (TS) approach of Kandasamy et al. [2018] and a purely random approach using Latin hypercube sampling (Random). The optimisation pipeline was constructed using GPyTorch [Gardner et al., 2018] and BoTorch [Balandat et al., 2020], and we have made this available online¹.

A zero-mean Gaussian process surrogate model with an isotropic Matérn 5/2 kernel, as recommended by Snoek et al. [2012] for modelling realistic functions, was used in all the experiments. Input variables were rescaled to reside in $[0, 1]^d$ and observations were rescaled to have zero-mean and unit variance. The models were initially trained on $M = 2d$ observations generated by maximin Latin hypercube sampling and then optimised for a further $200 - 2d$ function evaluations. Each optimisation run was repeated 51 times from a different set of Latin hypercube samples.

¹<http://www.github.com/georgedeath/aegis>

Table 1: Benchmark Functions and Dimensionality.

Name	d	Name	d
Branin	2	Ackley	5, 10
Eggholder	2	Michalewicz	5, 10
GoldsteinPrice	2	StyblinskiTang	5, 7, 10
SixHumpCamel	2	Hartmann6	6
Hartmann3	3	Rosenbrock	7, 10

Sets of initial locations were common across all methods to enable statistical comparison. At each iteration, before new locations were selected, the hyperparameters of the GP were optimised by maximising the log likelihood using a multi-restart strategy [Shahriari et al., 2016] with L-BFGS-B [Byrd et al., 1995] and 10 restarts (Algorithm 1, line 3).

All experiments were repeated for $q \in \{4, 8, 16\}$ asynchronous workers. We followed the procedure of Kandasamy et al. [2018] to sample an evaluation time for each task, thereby simulating the asynchronous setting. A half-normal distribution was used for this with a scale parameter of $\sqrt{\pi}/2$, giving a mean runtime of 1 [Alvi et al., 2019].

The KB, LP and PLAyBOOK methods all used the EI acquisition function, which was selected because of its popularity. The expected improvement, the sampled function from the GP posterior in both TS and AEGiS, and the posterior mean in AEGiS were all optimised using the typical strategy [Balandat et al., 2020] of uniformly sampling $1000d$ locations in \mathcal{X} , optimising the best 10 of these using L-BFGS-B, and selecting the best as the optimal location. In AEGiS and TS, 2000 Fourier features were used for the decoupled sampling method of Wilson et al. [2020]. The approximate Pareto set $\tilde{\mathcal{P}}$ in AEGiS was found using NSGA-II [Deb et al., 2002]; see supplementary material.

Here, we report performance in terms of the logarithm of the simple regret R_t , which is the difference between the true minimum value $f(\mathbf{x}^*)$ and the best seen function evaluation up to iteration t : $\log(R_t) = \log(\min\{f_1, \dots, f_t\} - f(\mathbf{x}^*))$.

4.1 SYNTHETIC EXPERIMENTS

The methods were evaluated on the 15 synthetic benchmark functions listed in Table 1.² Figure 1 shows four illustrative convergence plots for the Branin, SixHumpCamel, Ackley, and StyblinskiTang benchmark functions for $q \in \{4, 8, 16\}$ asynchronous workers. Convergence plots and tabulated results for all functions are available in the supplementary material. As can be seen from the four sets of convergence plots, AEGiS (pink) is almost always better than TS (orange) on all four problems and values of q . Contrastingly, TS is sometimes worse than random search (blue); this is

²Formulae for all functions can be found at: <http://www.sfu.ca/~ssurjano/optimization.html>.

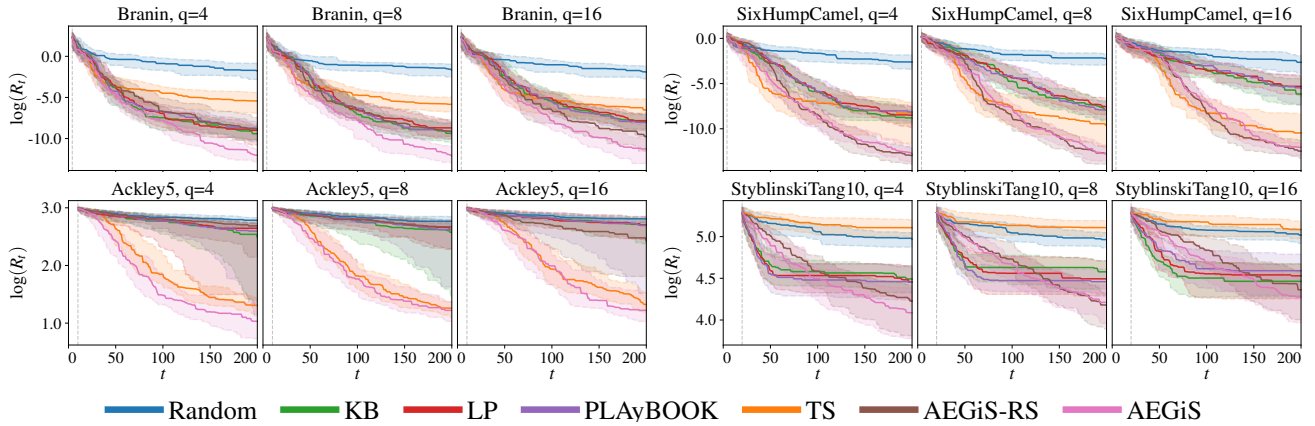


Figure 1: Illustrative convergence plots for four benchmark problems with $q \in \{4, 8, 16\}$ asynchronous workers. Each plot shows the median log simple regret, with shading representing the interquartile range over 51 runs.

particularly evident on the StyblinskiTang problem. This highlights the efficacy of both additional exploration and exploitation in AEGiS relative to the TS algorithm. The SixHumpCamel convergence plots show an interesting trend with respect to the penalisation-based methods, namely that they deteriorate at an accelerated rate in comparison to the TS-based methods. We suspect that this is because errors in estimating the correct degree of penalisation accumulate more noticeably as q increases and therefore more function evaluations are placed in unpromising regions.

Figure 2 summarises the performance of each of the evaluated methods for different numbers of workers (q). It shows the number of times each method is the best, i.e. has the lowest median regret over the 51 optimisation runs, or is statistically indistinguishable from the best method, according to a one-sided, paired Wilcoxon signed-rank test [Knowles et al., 2006] with Holm-Bonferroni correction [Holm, 1979] ($p \geq 0.05$). As is clear from the figure, AEGiS achieves a strong level of performance, and is the best (or equivalent to the best) on 10 out of the 15 functions. AEGiS-RS, which samples exploratory locations from the entire feasible space, is at best equivalent to AEGiS and generally worse. This indicates that using a more informative random sampling scheme, such as selecting from approximate Pareto set of the trade-off between exploitation ($\mu(\mathbf{x})$) and exploration ($\sigma^2(\mathbf{x})$), provides a meaningful improvement to performance. In contrast to both AEGiS and TS, which barely decrease in performance as the batch size q increases, the EI-based methods (PLAyBOOK, LP and KB) show a much larger reduction in relative performance.

4.2 HYPERPARAMETER OPTIMISATION

Like Souza et al. [2020], we optimise the hyperparameters of three meta-surrogate optimisation tasks corresponding to a SVM, a fully connected neural network (FC-NET) and XGBoost. These are part of the Profet benchmark [Klein

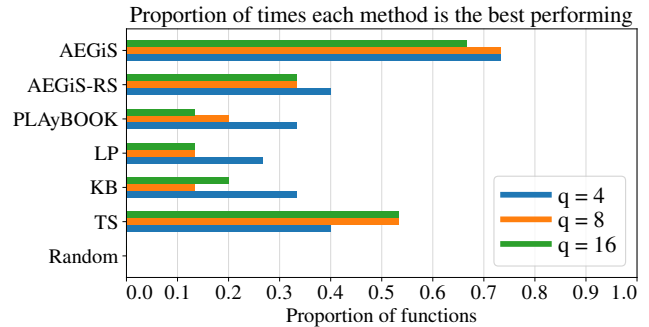


Figure 2: Synthetic function optimisation summary. Bar lengths correspond to the proportion of times that a method is best or statistically equivalent to the best method across the 15 synthetic functions, for $q \in \{4, 8, 16\}$ workers.

et al., 2019, Paleyes et al., 2019] and are drawn from generative models built using performance data on multiple datasets. They aim to mimic the landscape of expensive hyperparameter optimisation tasks, preserving the global landscape characteristics of the modelled methods and giving local variation between function draws from the generative model. This allows far more evaluations to be carried out than would be possible with the modelled problems.

Here, we optimise 51 instances of each meta-surrogate, repeating the optimisation $N = 20$ times per instance with different paired training data for each of the N runs. Details of the construction of each problem are given in the supplementary material. The SVM problem had 2 parameters; the FC-NET problem 6; and the XGBoost problem 8.

We compare each evaluated method by computing the average ranking score in every iteration for each problem instance. We follow Feurer et al. [2015] and compute this by, for each problem instance, drawing a bootstrap sample of 1000 runs out of the 7^N possible combinations and calculating the fractional ranking after each of the 200 iterations.

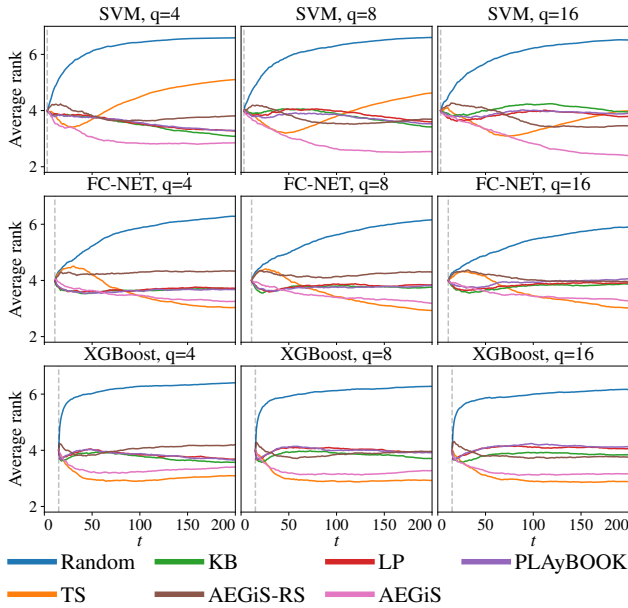


Figure 3: Average ranking scores for the three meta-surrogate benchmarks for $q \in \{4, 8, 16\}$.

The ranks are then averaged over all 51 problem instances.

Figure 3 shows the average ranks of the methods on all three problems for $q \in \{4, 8, 16\}$. Similarly to the synthetic functions, all three penalisation methods (KB, LB and PLAyBOOK) perform comparably and AEGiS is consistently superior to them after roughly 50 function evaluations. On the FC-NET and XGBoost problems TS performs similarly to AEGiS, and indeed is better on the XGBoost problem. However, AEGiS-RS performs consistently worse, supporting the contention that choosing random locations that lie in the Pareto set is beneficial.

4.3 ACTIVE LEARNING FOR ROBOT PUSHING

Following Wang and Jegelka [2017] and De Ath et al. [2021], we optimise the control parameters for two instance-based active learning robot pushing problems [Wang et al., 2018]. In push4, a robot pushes an object towards an unknown target location. It receives the object-target distance once it has finished pushing. The location of the robot, the orientation of its pushing hand and for how long it pushes are the parameters to be optimised with the goal of minimising this distance. The push8 problem has two robots that receive the distance between their respective objects and targets after pushing — we minimise the sum of these distances. In both problems the target location(s) are chosen randomly, with the constraint that they cannot overlap. We note that the push8 problem is much more difficult because the robots can block each other, and so each problem instance may not be completely solvable. Further details of the problems are provided in the supplementary material.

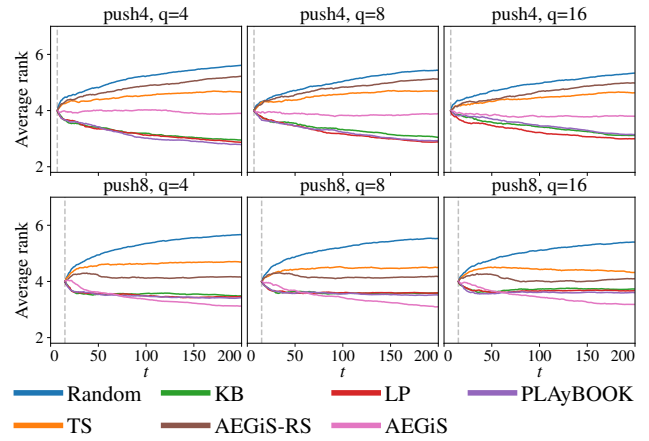


Figure 4: Average ranking scores for the two robot pushing problems for $q \in \{4, 8, 16\}$.

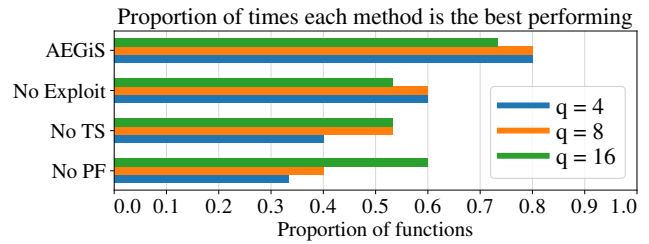


Figure 5: Ablation study summary. The bars correspond to the proportion of times that a method is best or statistically equivalent to the best method across the 15 synthetic functions, for $q \in \{4, 8, 16\}$.

Average rank score plots, calculated in the same way as described in Section 4.2, are shown in Figure 4. The EI-based methods continue to have very similar performance to one another, and are the best methods on push4, and, even though it is not the best method, AEGiS still outperforms TS consistently. On push8, the EI-based methods initially improve on AEGiS in terms of average rank but then stagnate and are overtaken by AEGiS after roughly 75 function evaluations.

4.4 ADDITIONAL EXPERIMENTS

Lastly, we describe an ablation study, experiments on setting the degree of deliberate exploration ($\epsilon = \epsilon_T + \epsilon_P$), together with the optimal ratio $\epsilon_T : \epsilon_P$. Finally, we compare AEGiS, EI and TS in the sequential BO setting ($q = 1$). See supplementary material for full results.

Ablation Study Here, we conduct an ablation study on AEGiS using the 15 synthetic benchmark functions used in Section 4.1. Since AEGiS-RS is consistently outperformed by AEGiS, we omit it from this study. We compare to AEGiS in the following ways: *No Exploit*, without the exploitation ($\epsilon_T = \epsilon_P = 1/2$); *No TS*, with no Thompson

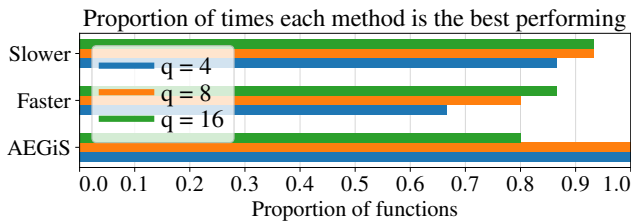


Figure 6: Selecting ϵ : *faster* and *slower* correspond to quicker and slower rates of ϵ decay than AEGiS, resulting in more or less exploitation respectively.

sampling ($\epsilon_T = 0; \epsilon_P = \epsilon$); *No PF*, without Pareto set selection ($\epsilon_T = \epsilon, \epsilon_P = 0$). In all cases $\epsilon = \min(2/\sqrt{d}, 1)$.

Figure 5 shows that all three components combine to give the best results and that removing any of the three results in an inferior algorithm. As can be seen, the removal of either TS or Pareto set selection considerably reduces the performance of the algorithm. Note that the results for *No Exploit* are inflated because the method is equivalent to AEGiS on the 5 out of the 15 benchmark functions that have $d \leq 4$ dimensions.

Setting ϵ We investigate the rate at which the degree of deliberate exploitation of the surrogate model’s posterior mean function should increase with problem dimensionality d . In AEGiS, exploitation is carried out $1 - (\epsilon_T + \epsilon_P)$ proportion of the time. As above we chose $\epsilon_T = \epsilon_P = \epsilon/2$ and ϵ was chosen *a priori* to decay like $1/\sqrt{d}$, i.e. $\epsilon = \min(2/\sqrt{d}, 1)$. Here, we bracket this decay rate with a quicker decay (more exploitation) and a slower decay (more exploration). Specifically, we evaluate AEGiS on the 15 synthetic test functions using $\epsilon = \min(2/(d - 2), 1)$ and $\epsilon = \min(2/\log(d + 3), 1)$ labelled *slower* and *faster* respectively. These decay functions were chosen because they match AEGiS and provide no exploitation when $d \leq 4$, i.e. $\epsilon = 0.5$. See the supplementary material for a visual comparison of the rates. Figure 6 summarises the results on the 15 test functions. When using a smaller number of workers ($q \in \{4, 8\}$) the $2/\sqrt{d}$ decay gives superior performance. However, when $q = 16$ there is little to differentiate between the three rates. We note that when d is large the faster rate of the decay (more exploitation for a given d) is superior; see the supplementary material for full details.

Proportion of TS to PF selection In the above evaluations TS and PF were selected with equal probability $\epsilon_T = \epsilon_P$. Here, we investigate the proportion of times that TS should be chosen over Pareto set selection. Specifically we evaluate AEGiS on the synthetic benchmark functions for $q \in \{4, 8, 16\}$ with the split between exploitation, TS and Pareto set selection being $1 - \epsilon$, $\epsilon_T = \gamma\epsilon$ and $\epsilon_P = (1 - \gamma)\epsilon$ respectively and $\gamma \in \{0, 0.1, \dots, 1\}$. Note that two of the methods evaluated in the earlier ablation study, *No TS* and *No PF*, correspond to $\gamma = 0$ and

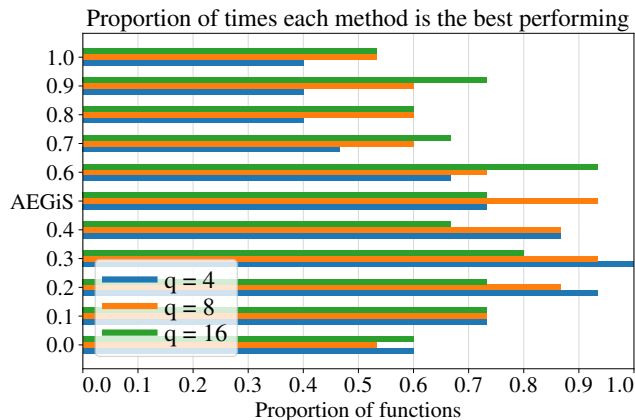


Figure 7: Proportion of TS to Pareto set selection for proportions $\gamma \in \{0, 0.1, \dots, 1\}$ (vertical axis). Note that the label AEGiS corresponds to $\gamma = 0.5$.

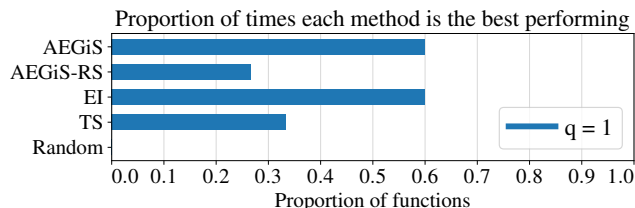


Figure 8: Synthetic function optimisation summary for the sequential setting ($q = 1$). Bar lengths correspond to the proportion of times that a method is best or statistically equivalent to the best method across the 15 synthetic functions.

$\gamma = 1$ respectively. Figure 7 summarises the results on the 15 synthetic benchmark functions. It shows that using more TS than Pareto set selection appears detrimental to AEGiS, whereas using less TS, i.e. a smaller value of γ leads to slightly improved average performance over 0.5. Interestingly, while AEGiS generally performed better with a smaller value of γ , for some test functions, e.g. Rosenbrock, larger values were preferred – see the supplementary material for all convergence plots and tabulated results. In general we recommend the standard setting $\epsilon_T = \epsilon_P = \epsilon/2$.

Sequential BO Finally, we evaluate AEGiS in the sequential setting, i.e. with $q = 1$. In this setting KB, LP, and PLAYBOOK are equivalent because they all select the next (and only) location to be evaluated using EI. Therefore, we compare AEGiS and AEGiS-RS to TS, EI, and Latin hypercube sampling. As shown in Figure 8, AEGiS and EI had approximately equivalent performance, and they were both superior to AEGiS-RS and TS. This again emphasises the importance of sampling from $\hat{\mathcal{P}}$. An ablation study was also carried out with similar findings: removing the sampling from $\hat{\mathcal{P}}$ led to a large reduction in performance. This was a larger drop in performance in comparison to the ablation study using larger values of q . We suspect that this is because the model will naturally be of a higher quality at

any iteration for $q = 1$ and thus selection from $\tilde{\mathcal{P}}$ will be even more informative to the optimisation process. Different rates of ϵ decay were also investigated as before, and similar results were found: both decreasing and increasing the rate of decay led to worse performance, with the largest decrease coming from increasing the rate and thus increasing exploitation further.

5 CONCLUSION

Practical optimisation of expensive functions can often make use of parallel hardware to rapidly obtain results. The AEGiS method makes best use of hardware resources through asynchronous Bayesian optimisation by combining greedy exploitation of the surrogate mean prediction with Thompson sampling and exploratory moves from the approximate Pareto set that maximally trades off exploration and exploitation. The ablation study verifies the importance of each of these components. With only a single worker this simple algorithm is no worse than BO using expected improvement. As the problem dimension increases deliberate exploratory moves are less necessary because of the inadvertent exploration due to modelling inaccuracies. We showed empirically on a wide range of problems that setting $\epsilon \propto 2/\sqrt{d}$ is more efficient than faster or slower rates of reducing the amount of exploration carried out as the problem dimensionality d increases.

Unlike other methods, such as PLAYBOOK and Thompson sampling, AEGiS cannot be trivially transformed into a synchronous batch BO method, and therefore it cannot be directly compared to a synchronous version of itself. The closest method in spirit is the ϵ -shotgun method of De Ath et al. [2020]. It scatters $q - 1$ samples around a central location, where the samples are distributed according the properties of the surrogate model’s landscape. The central location is chosen to be either the most exploitative location or a randomly selected location from the Pareto set. Interestingly the authors of both PLAYBOOK and Thompson sampling empirically found that, even though there is less information available for the selection of each asynchronous location, their methods outperformed their synchronous counterparts. We note that AEGiS is more effective than both PLAYBOOK and Thompson sampling and, therefore, also their synchronous equivalents.

Although AEGiS is robust to both the amount of exploration ($\epsilon_T + \epsilon_P$) and the ratio $\epsilon_T : \epsilon_P$, the combination of greedy exploitation, Thompson sampling and selection from the exploration-exploitation Pareto set is a challenge to obtaining non-trivial bounds on the convergence rate. While future work will concentrate on providing such theoretical guarantees, our extensive empirical evaluation shows that AEGiS is a simple, practical and robust method for asynchronous batch Bayesian optimisation.

Acknowledgements

This work was supported by Innovate UK, grant number 104400. The authors would like to acknowledge the use of the University of Exeter High-Performance Computing (HPC) facility.

References

- Ahsan Alvi, Binxin Ru, Jan-Peter Calliess, Stephen Roberts, and Michael A. Osborne. Asynchronous batch Bayesian optimisation with improved local penalisation. In *International Conference on Machine Learning*, pages 253–262. PMLR, 2019.
- Javad Azimi, Alan Fern, and Xiaoli Z Fern. Batch Bayesian optimization via simulation matching. In *Advances in Neural Information Processing Systems*, pages 109–117, 2010.
- Maximilian Balandat, Brian Karrer, Daniel R. Jiang, Samuel Daulton, Benjamin Letham, Andrew Gordon Wilson, and Eytan Bakshy. BoTorch: Programmable Bayesian optimization in PyTorch. In *Advances in Neural Information Processing Systems*, 2020.
- Eric Brochu, Vlad M. Cora, and Nando de Freitas. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning, 2010. *arXiv:1012.2599*.
- Adam D. Bull. Convergence rates of efficient global optimization algorithms. *Journal of Machine Learning Research*, 12(88):2879–2904, 2011.
- Richard H Byrd, Peihuang Lu, Jorge Nocedal, and Ciyou Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5):1190–1208, 1995.
- Daniele Calandriello, Luigi Carratino, Alessandro Lazaric, Michal Valko, and Lorenzo Rosasco. Gaussian process optimization with adaptive sketching: Scalable and no regret. In *Conference on Learning Theory*, pages 533–557. PMLR, 2019.
- Yutian Chen, Aja Huang, Ziyu Wang, Ioannis Antonoglou, Julian Schrittwieser, David Silver, and Nando de Freitas. Bayesian optimization in AlphaGo, 2018. *arXiv:1812.06855*.
- Erik A. Daxberger and Bryan Kian Hsiang Low. Distributed batch Gaussian process optimization. In *International Conference on Machine Learning*, pages 951–960. PMLR, 2017.
- George De Ath, Richard M. Everson, Jonathan E. Fieldsend, and Alma A. M. Rahat. ϵ -shotgun: ϵ -greedy batch

- Bayesian optimisation. In *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*, pages 787–795. Association for Computing Machinery, 2020.
- George De Ath, Richard M. Everson, Alma A. M. Rahat, and Jonathan E. Fieldsend. Greed is good: Exploration and exploitation trade-offs in Bayesian optimisation. *ACM Transactions on Evolutionary Learning and Optimization*, 1(1), 2021.
- Alessandro De Palma, Celestine Mendler-Dünner, Thomas Parnell, Andreea Anghel, and Haralampos Pozidis. Sampling acquisition functions for batch Bayesian optimization, 2019. *arXiv:1903.09434*.
- K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2): 182–197, 2002.
- Thomas Desautels, Andreas Krause, and Joel W. Burdick. Parallelizing exploration-exploitation tradeoffs in Gaussian process bandit optimization. *Journal of Machine Learning Research*, 15(119):4053–4103, 2014.
- Matthias Feurer, Jost Tobias Springenberg, and Frank Hutter. Initializing Bayesian hyperparameter optimization via meta-learning. In *AAAI Conference on Artificial Intelligence*, pages 1128–1135. AAAI Press, 2015.
- Jacob Gardner, Geoff Pleiss, Kilian Q Weinberger, David Bindel, and Andrew G Wilson. GPyTorch: Blackbox matrix-matrix Gaussian process inference with GPU acceleration. In *Advances in Neural Information Processing Systems*, pages 7576–7586. 2018.
- David Ginsbourger, Rodolphe Le Riche, and Laurent Carraro. A multi-points criterion for deterministic parallel global optimization based on Gaussian processes, 2008. <https://hal.archives-ouvertes.fr/hal-00260579>.
- David Ginsbourger, Rodolphe Le Riche, and Laurent Carraro. Kriging is well-suited to parallelize optimization. In *Computational Intelligence in Expensive Optimization Problems*, volume 2, pages 131–162. Springer, 2010.
- Javier González, Zhenwen Dai, Philipp Hennig, and Neil Lawrence. Batch Bayesian optimization via local penalization. In *Artificial Intelligence and Statistics*, pages 648–657. PMLR, 2016.
- Sunil Gupta, Alistair Shilton, Santu Rana, and Svetha Venkatesh. Exploiting strategy-space diversity for batch Bayesian optimization. In *International Conference on Artificial Intelligence and Statistics*, pages 538–547. PMLR, 2018.
- Philipp Hennig and Christian J. Schuler. Entropy search for information-efficient global optimization. *Journal of Machine Learning Research*, 13(1):1809–1837, 2012.
- José Miguel Hernández-Lobato, Matthew W. Hoffman, and Zoubin Ghahramani. Predictive entropy search for efficient global optimization of black-box functions. In *Advances in Neural Information Processing Systems*, pages 918–926. MIT Press, 2014.
- José Miguel Hernández-Lobato, James Requeima, Edward O. Pyzer-Knapp, and Alán Aspuru-Guzik. Parallel and distributed Thompson sampling for large-scale accelerated exploration of chemical space. In *International Conference on Machine Learning*, pages 1470–1479. PMLR, 2017.
- Sture Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979.
- Janis Janusevskis, Rodolphe Le Riche, David Ginsbourger, and Ramunas Girdziusas. Expected improvements for the asynchronous parallel global optimization of expensive functions. In *Learning and Intelligent Optimization*, volume 7219, pages 413–418. Springer, 2012.
- Mikkel T. Jensen. Reducing the run-time complexity of multiobjective EAs: The NSGA-II and other algorithms. *IEEE Transactions on Evolutionary Computation*, 7(5): 503–515, 2003.
- Donald R. Jones, Matthias Schonlau, and William J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492, 1998.
- Kirthevasan Kandasamy, Akshay Krishnamurthy, Jeff Schneider, and Barnabas Poczos. Parallelised Bayesian optimisation via Thompson sampling. In *International Conference on Artificial Intelligence and Statistics*, pages 133–142. PMLR, 2018.
- Aaron Klein, Zhenwen Dai, Frank Hutter, Neil Lawrence, and Javier González. Meta-surrogate benchmarking for hyperparameter optimization. In *Advances in Neural Information Processing Systems*, pages 6270–6280. 2019.
- Joshua D. Knowles, Lothar Thiele, and Eckart Zitzler. A tutorial on the performance assesment of stochastic multiobjective optimizers. Technical Report TIK214, Computer Engineering and Networks Laboratory, ETH Zurich, Zurich, Switzerland, 2006.
- Harold J. Kushner. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *Journal Basic Engineering*, 86(1):97–106, 1964.
- Micheal D. McKay, Richard J. Beckman, and William J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979.

- Andrei Paleyev, Mark Pullin, Maren Mahsereci, Neil Lawrence, and Javier González. Emulation of physical processes with Emukit. In *Second Workshop on Machine Learning and the Physical Sciences, NeurIPS*, 2019.
- Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. In *Advances in Neural Information Processing Systems*, pages 1177–1184. 2008.
- Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian processes for machine learning*. The MIT Press, Boston, MA, 2006. ISBN 0-262-18253-X.
- Binxin Ru, Michael A. Osborne, Mark Mcleod, and Diego Granzio. Fast information-theoretic Bayesian optimisation. In *International Conference on Machine Learning*, pages 4384–4392. PMLR, 2018.
- Warren Scott, Peter Frazier, and Warren Powell. The correlated knowledge gradient for simulation optimization of continuous parameters using Gaussian process regression. *SIAM Journal on Optimization*, 21(3):996–1026, 2011.
- Amar Shah and Zoubin Ghahramani. Parallel predictive entropy search for batch global optimization of expensive objective functions. In *Advances in Neural Information Processing Systems*, pages 3330–3338. 2015.
- Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P. Adams, and Nando de Freitas. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE*, 104(1):148–175, 2016.
- Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical Bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems*, pages 2951–2959, 2012.
- Artur Souza, Luigi Nardi, Leonardo B. Oliveira, Kunle Olukotun, Marius Lindauer, and Frank Hutter. Prior-guided Bayesian optimization, 2020. *arXiv:2006.14608*.
- Niranjan Srinivas, Andreas Krause, Sham Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. In *International Conference on Machine Learning*, pages 1015–1022. Omnipress, 2010.
- William R. Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3/4):285–294, 1933.
- Zi Wang and Stefanie Jegelka. Max-value entropy search for efficient Bayesian optimization. In *International Conference on Machine Learning*, pages 3627–3635. PMLR, 2017.
- Zi Wang, Caelan Reed Garrett, Leslie Pack Kaelbling, and Tomas Lozano-Pérez. Active model learning and diverse action sampling for task and motion planning. In *International Conference on Intelligent Robots and Systems (IROS)*, pages 4107–4114. IEEE, 2018.
- James Wilson, Viacheslav Borovitskiy, Alexander Terenin, Peter Mostowsky, and Marc Deisenroth. Efficiently sampling functions from Gaussian process posteriors. In *International Conference on Machine Learning*, pages 10292–10302. PMLR, 2020.
- Jian Wu and Peter Frazier. The parallel knowledge gradient method for batch Bayesian optimization. In *Advances in Neural Information Processing Systems*, pages 3126–3134. 2016.