Bayesian Experimental Design Literature Review

Connections with Bayesian Active Learning, Bayesian Optimisation and Bayesian Reinforcement Learning

Adam Foster
Department of Statistics, University of Oxford

July 2021

1 Introduction

If true knowledge arises from empirical observations, it is natural to ask which kinds of observations we should actively seek out to further our understanding of nature. In its broadest sense, this is the question that the design of experiments seeks to answer. An *experimental design* is an allocation of resources—e.g. time, human attention, chemical reagents, physical space—that will be used to obtain empirical observations. The *design space* is the set of designs that we could feasibly choose for the experiment; the problem of experimental design is to pick a design to use for the real experiment. The choice of design is an important one: we could easily waste resources on poorly designed experiments that do not further our understanding. By carefully designing experiments, we can efficiently gather empirical observations that lead to new ideas, hypotheses, conclusions and models.

It is therefore unsurprising that we find experimental design to be a key concern in scientific disciplines as diverse as psychology (Myung et al., 2013), bioinformatics (Vanlier et al., 2012), pharmacology (Lyu et al., 2019), physics (Dushenko et al., 2020), neuroscience (Shababo et al., 2013), astronomy (Loredo, 2004) and engineering (Papadimitriou, 2004). It is also is a natural abstraction for several central problems in machine learning, including active learning (Houlsby et al., 2011; Gal et al., 2017), Bayesian optimisation (Hernández-Lobato et al., 2014; Shahriari et al., 2015) and exploration (Sun et al., 2011; Shyam et al., 2019).

In many practical cases, experimental design is not used just once. Indeed, many experiments are naturally *adaptive*: they are an iterative process in which we can select the designs for later iterations on the basis of data already gathered. This allows feedback from the outcome of one experiment iteration to be used to guide the design of the next iteration. This setting can be particularly powerful because, as we gain some information about the system, it may become clearer how we should proceed to design our experiments to investigate further, thereby honing in quickly on the truth.

To choose between different possible experimental designs requires an objective function. In general, the objective depends not only on known quantities (such as the cost of the experiment), but also on the not-yet-observed outcome of the experiment and potentially on other unobserved quantities. For example, the objective function for a chemical experiment might reward correctly synthesising a product, something that will only be observed once the experiment is completed. To reason about objective functions that depend on unknowns in this way requires the incorporation of some *a priori* knowledge. This *a priori* knowledge is then used to select the design before commencing the experiment. In this work, we focus on the Bayesian approach to this problem (Lindley, 1956, 1972; Chaloner and Verdinelli, 1995; Ryan et al., 2016; Foster et al., 2019) in which *a priori* knowledge is encoded in two ways—first, the specification of a model for the experiment, and second in the prior distribution for the unknown parameters of that model. Typically, the model itself is assumed to be correct. The prior distribution explicitly represents initial beliefs about unknown parameters of the model. Furthermore, uncertainty in the prior is exactly the epistemic uncertainty that can be reduced by running experiments and collecting data, resulting in more precise *a posteriori* knowledge.

In this literature review, we begin with a brief survey of foundational concepts in Bayesian data analysis

(Sec. 2). We then turn to the core theory of Bayesian experimental design (Sec. 3), discussing criteria that have been used within the statistics community, with an emphasis on expected information gain. In Sec. 4, we discuss computational methods for Bayesian experimental that have been used within statistics, and in Sec. 5 we discuss active learning. We then discuss models in which the target of experimental design is embedded in a larger model (Sec. 6); Bayesian optimisation (Sec. 7) is a specific instance of this. Finally, we delve into the theory of the sequential experimental design problem (Sec. 8), and highlight connections with exploration and Bayesian reinforcement learning (Sec. 9).

2 Background on Bayesian statistics

We first introduce necessary notation and key concepts in Bayesian data analysis[1]. The first ingredient of any Bayesian analysis is a full probability model that places a joint distribution over all observable and unobservable quantities. We denote the parameters of interest, also called the latent variable of interest, by $\theta \in \Theta$. This may be a scalar, vector, or a function depending on the model. We denote the observed data, or outcome, as $y \in \mathcal{Y}$. The full probability model is simply a probability distribution $p(\theta, y)$ on $\Theta \times \mathcal{Y}$. Typically, the full probability model can be factorised as

$$p(\theta, y) = p(\theta)p(y|\theta) \tag{1}$$

where $p(\theta)$ denotes the *prior* on θ, and $p(y|\theta)$ is the *likelihood* function[2], or sampling distribution.

Since we are interested in experimental design, we also introduce the *design* or *covariate* $\xi \in \Xi$. This is not typically treated as a random variable, because it is assumed to be directly under the experimenter's control. Instead, for each possible design ξ, we have a different probability model $p(\theta, y|\xi)$. Different choices of ξ should not alter our prior $p(\theta)$, thus the change in the probability model is only felt through the likelihood, so we can write $p(\theta, y|\xi) = p(\theta)p(y|\theta, \xi)$. Intuitively, this says that the design of the experiment ξ does not change the natural environment, but it can change the outcome of an experiment that we choose to run.

Once we have chosen ξ and run our experiment to obtain y, we can make probability statements about θ by applying Bayes' Rule to calculate the posterior

$$p(\theta|\xi, y) = \frac{p(\theta)p(y|\theta, \xi)}{\int_\Theta p(\theta')p(y|\theta', \xi)d\theta'} = \frac{p(\theta)p(y|\theta, \xi)}{p(y|\xi)}. \tag{2}$$

In general, actually performing Bayesian inference to calculate $p(\theta|\xi, y)$ can be computationally challenging.

2.1 Explicit and implicit models

If the likelihood $p(y|\theta, \xi)$ is known in closed form, then the probability model is called an *explicit likelihood* model. Most Bayesian statistics assumes an explicit likelihood. If no closed form likelihood is available, the model is an *implicit likelihood* model (Sisson et al., 2018). Implicit models often arise when θ, y and ξ are related by a simulator (Alsing et al., 2019; Brehmer et al., 2018; Gonçalves et al., 2020) that can produce samples of $p(y|\theta, \xi)$, but does not have a closed form probability density.

Similarly, if $p(\theta)$ is known in closed form, then the model is said to have an *explicit prior*, otherwise, the prior is said to be *implicit*.

2.2 Sequential data collection

So far, we have considered choosing ξ, collecting y, and analysing the data by computing $p(\theta|\xi, y)$. A more realistic setting is to consider a sequence $\xi_1 \dots \xi_T$ of designs with corresponding outcomes $y_1, \dots, y_T$. This

[1] More details on Bayesian data analysis can be found in modern textbooks on the topic, such as Gelman et al. (2013) and Kruschke (2014).

[2] Strictly, the likelihood describes the sampling distribution $p(y|\theta)$ as a function of θ for a fixed y; we use likelihood in a slightly looser sense to refer to $p(y|\theta)$ in general.

means that we run T different experiments with T different designs, each with its own corresponding outcome. The value of θ, although unknown, is assumed to be the same across all the T experiments—that means that we are conducting multiple experiments in the same natural environment to gather further information about it, instead of starting afresh in a new environment for each new experiment.

In an *exchangeable* model (Bloem-Reddy and Teh, 2019), the order of the experiments does not matter. This is equivalent (Øksendal, 2003) to the following factorisation of the full probability model

$$p(\theta, y_{1:T}|\xi_{1:T}) = p(\theta)\prod_{t=1}^{T} p(y_t|\theta, \xi_t). \tag{3}$$

for some random variable θ. The question is whether we can identify this θ with the model parameters of interest θ. In general, this is valid when there are no other model parameters besides θ. Indeed, in a full statistical model with parameters θ (Cox, 2006), it is common to assume that the outcomes of different experiments are independent given θ, which is equivalent to the factorisation in equation (3). We discuss the case in which there are other model parameters aside from θ in Sec. 6.

In non-exchangeable models, there is no assumption of conditional independence between experiments. Such models are uncommon, but can arise in settings such as time series (Pole et al., 2018). In a non-exchangeable model, the distribution of y_t can, for example, be influenced by y_{t-1} as well as by θ and ξ_t. Without loss of generality, the probability model for a non-exchangeable model can be written

$$p(\theta, y_{1:T}|\xi_{1:T}) = p(\theta)\prod_{t=1}^{T} p(y_t|\theta, \xi_{1:t}, y_{1:t-1}). \tag{4}$$

which encodes only the assumption that future experiments cannot affect the outcome of earlier experiments.

Static and adaptive experiments An orthogonal distinction in sequential experiments is how the designs are generated. In a *static* experiment, also called fixed, batch, or open loop (DiStefano III et al., 2014), the designs $\xi_1, \ldots, \xi_T$ are chosen before the beginning of the experiment. In an *adaptive* experiment (Myung et al., 2013), each ξ_t is chosen depending on data already seen $\xi_1, \ldots, \xi_{t-1}, y_1, \ldots, y_{t-1}$. A simple consequence of the likelihood principle (Barnard et al., 1962; Birnbaum, 1962) is that the mode in which the ξ_t are generated does not affect the posterior distribution on θ calculated from the data. Indeed, suppose each new design is chosen adaptively from a density $p(\xi_t|\xi_{1:t-1}, y_{1:t-1})$. Then the resulting posterior distribution is

$$p(\theta|\xi_{1:T}, y_{1:T}) = \frac{p(\theta)\prod_{t=1}^{T} p(\xi_t|\xi_{1:t-1}, y_{1:t-1})p(y_t|\theta, \xi_{1:t}, y_{1:t-1})}{\int_\Theta p(\theta')\prod_{t=1}^{T} p(\xi_t|\xi_{1:t-1}, y_{1:t-1})p(y_t|\theta', \xi_{1:t}, y_{1:t-1})d\theta'} \tag{5}$$

$$= \frac{\prod_{t=1}^{T} p(\xi_t|\xi_{1:t-1}, y_{1:t-1})\ p(\theta)\prod_{t=1}^{T} p(y_t|\theta, \xi_{1:t}, y_{1:t-1})}{\prod_{t=1}^{T} p(\xi_t|\xi_{1:t-1}, y_{1:t-1})\ \int_\Theta p(\theta')\prod_{t=1}^{T} p(y_t|\theta', \xi_{1:t}, y_{1:t-1})d\theta'} \tag{6}$$

$$= \frac{p(\theta)\prod_{t=1}^{T} p(y_t|\theta, \xi_{1:t}, y_{1:t-1})}{\int_\Theta p(\theta')\prod_{t=1}^{T} p(y_t|\theta', \xi_{1:t}, y_{1:t-1})d\theta'}, \tag{7}$$

which is independent of the mechanism of choosing designs.

2.3 Bayesian decision making

After collecting data $\xi_{1:T}, y_{1:T}$, suppose that we must choose some decision δ, for example whether to prescribe a medication or not. The Bayesian approach to selecting the optimal decision (Lindley, 1972; Robert, 2007) is to specify a utility function $U(\delta, \theta)$ which should assign a value to the decision δ *in the case that* θ is the true value of the unobserved parameter. The optimal decision is then found by maximising expected utility under the current posterior

$$\delta^* = \underset{\delta\in\Delta}{\arg\max} \quad \mathbb{E}_{p(\theta|\xi_{1:T}, y_{1:T})}[U(\delta, \theta)] \tag{8}$$

For a more extensive discussion of Bayesian decision theory, see Berger (2013).

3 Bayesian Experimental Design

Experimental design with a Bayesian data analysis model means choosing the design using the likelihood model and the prior $p(\theta)$ as *a priori* information. What criterion should be used to select the design? Following from Bayesian decision theory, Lindley (1972) proposed a decision-theoretic approach to Bayesian experimental design that focuses on maximising a utility. Chaloner and Verdinelli (1995) provides a more recent summary of Lindley's approach.

First, let us restrict ourselves to a single design ξ with outcome y, leaving the sequential design problem to Sec. 8. We augment the utility function of Sec. 2.3 to a utility $U(\delta, \theta, \xi, y)$ that incorporates potential costs of the experimental design and outcome. Whilst our discussion in Sec. 2.3 assumed that the data ξ, y had already been gathered, we now need to consider the choice of the design ξ as well as the decision δ. The order of operation for the experimenter is as follows:

1. choose design ξ;
2. perform experiment with design ξ, obtaining experimental outcome y;
3. compute the posterior $p(\theta|\xi, y)$;
4. choose the design δ to maximise $\mathbb{E}_{p(\theta|\xi,y)}[U(\delta, \theta, \xi, y)]$.

In order to choose ξ optimally, we should therefore consider the different possible observations y that could arise. Specifically, we will choose ξ to maximise the expected utility, taking an outer expectation over the observation y using the Bayesian marginal (also called prior predictive) distribution $p(y|\xi) = \mathbb{E}_{p(\theta)}[p(y|\theta, \xi)]$. This leads to the following method of choosing the optimal design

$$\xi^* = \arg\max_{\xi \in \Xi} \ \mathbb{E}_{p(y|\xi)} \left[\max_{\delta \in \Delta} \ \mathbb{E}_{p(\theta|\xi,y)}[U(\delta, \theta, \xi, y)] \right]. \tag{9}$$

Proposition 1 (Lindley (1972))**.** *It is not necessary to introduce randomness into the selection of* ξ*.*

Proof. Suppose we consider a randomised way of selecting ξ with distribution $p(\xi)$. The expected reward of this approach is

$$\mathbb{E}_{p(\xi)p(y|\xi)} \left[\max_{\delta \in \Delta} \ \mathbb{E}_{p(\theta|\xi,y)}[U(\delta, \theta, \xi, y)] \right] \leq \sup_{\xi \in \Xi} \mathbb{E}_{p(y|\xi)} \left[\max_{\delta \in \Delta} \ \mathbb{E}_{p(\theta|\xi,y)}[U(\delta, \theta, \xi, y)] \right] \tag{10}$$

where the righthand side is the expected utility using the non-random ξ^*. So a randomised design is not required. □

The remaining piece of the puzzle is to select a utility function. Some applications feature a highly problem-specific utility. In other cases, we can rely on general purpose utilities.

3.1 Expected Information Gain

Perhaps the most well-studied of all criteria for Bayesian experimental design is expected information gain (EIG). Within Bayesian experimental design, EIG appears to be dominant in a number of fields. EIG was proposed by Lindley (1956). Recent statistical review papers (Chaloner and Verdinelli, 1995; Ryan et al., 2016) give EIG pride of place within Bayesian experimental design. In psychology, Myung et al. (2013) promote the use of EIG to run adaptive trials. Several toolboxes (Watson, 2017; Vincent and Rainforth, 2017) have been designed specifically for the problem of performing adaptive psychology trials using EIG as the criterion for selecting designs. Heck and Erdfelder (2019) suggest EIG for experimental design for cognitive models and Cavagnaro et al. (2010) consider its application in the context of model discrimination in cognitive science. Shababo et al. (2013) applied EIG maximisation within a Bayesian model of neural microcircuits to choose the right subset of neurons to stimulate in an experiment. Dushenko et al. (2020)

proposed EIG as a criterion for designing measurement settings in magnetometry. In biochemistry, Busetto et al. (2009) compared EIG with several other criteria for the design of experiments for biochemical dynamical systems, finding EIG to perform best. In pharmacology, Lyu et al. (2019); Foster et al. (2020) applied EIG maximisation to design experiments to calibrate a docking model. Loredo (2004) used EIG for active exploration, specifically investigating the scheduling of observations of a star to characterise the orbit of a planet. EIG has also been used in active learning, Bayesian optimisation and reinforcement learning. We discuss these fields separately in Sections 5, 7 and 9.

There are several reasons for the dominance of the EIG. First, it has mathematical properties that make it very natural for describing information gained from experimentation. We discuss some key properties of EIG in this section, and we discuss EIG in sequential settings in Sec. 8. More practically, EIG applies to a range of linear and nonlinear models (unlike some criteria which are more restricted in their applicability) and handles both continuous and discrete θ.

What does EIG measure? EIG quantifies the amount of information that the experiment with design ξ is expected to produce about the unknown parameter of interest θ. A higher EIG indicates that doing the experiment with design ξ is likely to produce data that will be helpful in reducing uncertainty about the true value of θ.

To precisely define EIG, we utilise the rigorous probabilistic definition of information that was first given by Shannon (1948). Lindley (1956) used this work to quantify the information provided by an experiment. Lindley began by considering the Shannon *entropy* of a random variable θ

$$H[p(\theta)] = -\mathbb{E}_{p(\theta)}[\log p(\theta)]. \tag{11}$$

One interpretation of entropy is uncertainty in what the true value of θ is. In the experimental design context, we measure the amount of information that is gained about θ by performing the experiment with design ξ and obtaining outcome y using the reduction in entropy from the prior to the posterior. This is referred to as the information gain (IG)

$$U_{\mathcal{I}}(\xi, y) = \mathbb{E}_{p(\theta|\xi,y)}[\log p(\theta|\xi, y)] - \mathbb{E}_{p(\theta)}[\log p(\theta)]. \tag{12}$$

To obtain an objective function for ξ, we can use this utility within the decision-theoretic framework laid out in the preceding section. We substitute $U_{\mathcal{I}}$ into equation (9), noting that this utility does not involve any additional decision δ. This gives the overall objective function to select ξ: the *expected information gain* (EIG), formed by taking the expectation of $U_{\mathcal{I}}$ over $p(y|\xi)$, giving

$$\mathcal{I}(\xi) = \mathbb{E}_{p(y|\xi)}\left[\mathbb{E}_{p(\theta|\xi,y)}[\log p(\theta|\xi, y)] - \mathbb{E}_{p(\theta)}[\log p(\theta)]\right]. \tag{13}$$

Proposition 2 (Lindley (1956))**.** *EIG is the mutual information between y and θ under design ξ.*

Proof. By repeatedly using Bayes Theorem, we have

$$\begin{aligned}
\mathcal{I}(\xi) &= \mathbb{E}_{p(y|\xi)}\left[\mathbb{E}_{p(\theta|\xi,y)}[\log p(\theta|\xi, y)] - \mathbb{E}_{p(\theta)}[\log p(\theta)]\right] && (14)\\
&= \mathbb{E}_{p(y|\xi)p(\theta|\xi,y)}[\log p(\theta|\xi, y)] - \mathbb{E}_{p(\theta)}[\log p(\theta)] && (15)\\
&= \mathbb{E}_{p(\theta)p(y|\theta,\xi)}[\log p(\theta|\xi, y) - \log p(\theta)] && (16)\\
&= \mathbb{E}_{p(\theta)p(y|\theta,\xi)}\left[\log \frac{p(\theta|\xi, y)}{p(\theta)}\right] && (17)\\
&= \mathbb{E}_{p(\theta)p(y|\theta,\xi)}\left[\log \frac{p(\theta)p(y|\theta, \xi)}{p(\theta)p(y|\xi)}\right]. && (18)
\end{aligned}$$

□

Proposition 3. *EIG is unchanged under invertible reparametrisations of θ and y.*

Proof. This follows from the well-known property of mutual information (Cover, 1999). □

Proposition 4 (Bernardo (1979))**.** *EIG can equivalently be derived from the KL-divergence utility*

$$U_{KL}(\xi, y) = \mathrm{KL}\left(p(\theta|\xi, y) \| p(\theta)\right). \tag{19}$$

Proof. Substituting this utility into equation (9) gives us

$$\begin{aligned} I_{KL}(\xi) &= \mathbb{E}_{p(y|\xi)}\left[\mathrm{KL}\left(p(\theta|\xi, y) \| p(\theta)\right)\right] && (20)\\ &= \mathbb{E}_{p(y|\xi)p(\theta|y,\xi)}\left[\log \frac{p(\theta|y,\xi)}{p(\theta)}\right] && (21)\\ &= \mathbb{E}_{p(\theta)p(y|\theta,\xi)}\left[\log \frac{p(\theta|y,\xi)}{p(\theta)}\right] = \mathcal{I}(\xi) \text{ by equation (17).} && (22) \end{aligned}$$

□

Proposition 5 (Lindley (1956))**.** *EIG is convex in the likelihood.*

Proof. Let $\lambda \in [0, 1]$ and ξ_0, ξ_1 be two designs. Suppose there exists a design ξ_λ with the following likelihood

$$p(y|\theta, \xi_\lambda) = \lambda p(y|\theta, \xi_0) + (1 - \lambda)p(y|\theta, \xi_1). \tag{23}$$

Then Lindley (1956) Theorem 6 showed that

$$\mathcal{I}(\xi_\lambda) \leq \lambda \mathcal{I}(\xi_0) + (1 - \lambda)\mathcal{I}(\xi_1). \tag{24}$$

□

Proposition 6 (Sebastiani and Wynn (2000))**.** *EIG can be written as* $\mathcal{I}(\xi) = \mathbb{E}_{p(\theta)}\left[H[p(y|\xi)] - H[p(y|\theta, \xi)]\right]$*. Furthermore, when* $H[p(y|\theta, \xi)]$ *does not depend on* ξ*, EIG maximisation is equivalent to maximum entropy design which selects* ξ *to maximise* $H[p(y|\xi)]$*.*

Proof. Starting from Proposition 2, we have

$$\begin{aligned} \mathcal{I}(\xi) &= \mathbb{E}_{p(\theta)p(y|\theta,\xi)}\left[\log \frac{p(\theta)p(y|\theta,\xi)}{p(\theta)p(y|\xi)}\right] && (25)\\ &= \mathbb{E}_{p(\theta)p(y|\theta,\xi)}[\log p(y|\theta, \xi) - \log p(y|\xi)] && (26)\\ &= \mathbb{E}_{p(\theta)p(y|\theta,\xi)}[\log p(y|\theta, \xi)] - \mathbb{E}_{p(y|\xi)}[\log p(y|\xi)] && (27)\\ &= \mathbb{E}_{p(\theta)}\left[H[p(y|\xi)] - H[p(y|\theta, \xi)]\right]. && (28) \end{aligned}$$

Now, if $H[p(y|\theta, \xi)]$ is independent of ξ, then we have $\mathcal{I}(\xi) = H[p(y|\xi)] + \text{const.}$, so EIG maximisation and maximum entropy design lead to the same optimal design. □

Remark 7 (Smith and Gal (2018))**.** *EIG can be interpreted as a measure of epistemic uncertainty.*

Proof. Equation (28) breaks the EIG into two terms. The first is the total entropy $H[p(y|\xi)]$, called the predictive entropy. The second is $-\mathbb{E}_{p(\theta)}[H[p(y|\theta, \xi)]]$, which represents the expectation of the uncertainty in y *conditional on* θ. We can view this as a measure of aleatoric uncertainty—uncertainty which cannot be eliminated by knowing θ exactly. The EIG is the difference between the total and aleatoric uncertainties, hence we can interpret it as epistemic uncertainty—the part of $H[p(y|\xi)]$ that can be reduced by learning about θ. □

This interpretation does have its limitations. First, this definition of epistemic uncertainty is a model-dependent quantity—if we choose a more powerful model, then some variation that had previously been characterised as aleatoric would now be seen as epistemic. Second, the interpretation does not hold true in the case that θ is a function of a larger set of model parameters ψ, as in Sec. 6. This is because the term $-\mathbb{E}_{p(\theta)}[H[p(y|\theta, \xi)]]$ no longer represents aleatoric uncertainty, as it also includes some uncertainty that arises from not knowing the value of ψ.

Other important features of the EIG in sequential experiments will be discussed in Section 8.

3.2 Alphabetic criteria

3.2.1 Non-Bayesian linear model

The alphabetic criteria were initially proposed in the context of non-Bayesian experimental design for the linear model

$$y|\theta, \xi \sim N(\xi\theta, \sigma^2) \tag{29}$$

where ξ is the $n \times p$ design matrix and θ is a p-vector. (In a linear model, we would conventionally replace $\theta \to \beta$ and $\xi \to X$.) The least squares estimator for θ is $\hat{\theta} = (\xi^\top \xi)^{-1}\xi^\top y$. In frequentist analysis of this estimator, the covariance matrix of $\hat{\theta}$ is proportional to $(\xi^\top \xi)^{-1}$. To guide the choice of ξ, Box (1982) discussed the following notions of optimality of ξ

A-optimality minimise $\mathrm{Tr}(\xi^\top \xi)^{-1}$, or more generally, minimise $\mathrm{Tr}\, A(\xi^\top \xi)^{-1}$ for a matrix A;

D-optimality minimise $\det(\xi^\top \xi)^{-1}$;

E-optimality minimise $\max_i \lambda_i$, where $\lambda_1, \ldots, \lambda_p$ are the eigenvalues of $(\xi^\top \xi)^{-1}$;

G-optimality minimise $\sup_{c \in \mathcal{C}} c^\top (\xi^\top \xi)^{-1} c$, where $\mathcal{C}$ is some target region for prediction.

Other alphabetic criteria include

c-optimality (Elfving, 1952) minimise $c^\top (\xi^\top \xi)^{-1} c$ for some vector c.

T-optimality (Atkinson and Fedorov, 1975) for model discrimination, which maximises the minimal deviation between a null model and an alternative.

Several key results relate these classical criteria, such as Kiefer and Wolfowitz (1959).

3.2.2 Bayesian linear model

The alphabetic criteria can be extended to Bayesian linear models (Chaloner and Verdinelli, 1995), using the observation that the posterior covariance matrix for θ is proportional to $(\xi^\top \xi + \Sigma_0^{-1})^{-1}$ when we augment the model in equation (29) with a Gaussian prior $\theta \sim N(0, \Sigma_0)$. This allows a direct generalisation of the alphabetic criteria with $(\xi^\top \xi + \Sigma_0^{-1})^{-1}$ playing the role of $(\xi^\top \xi)^{-1}$. For example, Bayesian A-optimality minimises $\mathrm{Tr}(\xi^\top \xi + \Sigma_0^{-1})^{-1}$, and Bayesian D-optimality minimises $\det(\xi^\top \xi + \Sigma_0^{-1})^{-1}$.

Proposition 8 (Chaloner and Verdinelli (1995))**.** *For a Bayesian linear model, Bayesian D-optimality and EIG optimality are equivalent.*

Proof. In the Bayesian linear model, the posterior on θ is Gaussian with covariance matrix that is proportional to $(\xi^\top \xi + \Sigma_0^{-1})^{-1}$, and is independent of y. The entropy of this Gaussian posterior is $\frac{1}{2} \log \det(\xi^\top \xi + \Sigma_0^{-1})^{-1}$ + a constant. Substituting this into equation (13), the EIG for the Bayesian linear model is

$$\mathcal{I}(\xi) = \tfrac{1}{2} \log \det \Sigma_0 - \tfrac{1}{2} \log \det(\xi^\top \xi + \Sigma_0^{-1})^{-1} - \text{const} = -\tfrac{1}{2} \log \det(\xi^\top \xi + \Sigma_0^{-1})^{-1} + \text{const}'. \tag{30}$$

Thus, EIG optimality (maximise $\mathcal{I}(\xi)$) and Bayesian D-optimality (minimise $\det(\xi^\top \xi + \Sigma_0^{-1})^{-1}$) lead to the same optimal design. □

3.2.3 Bayesian non-linear models

The 'classical' approach (Tsutakawa, 1972; Chaloner and Verdinelli, 1995) to generalising the alphabetic criteria to non-linear Bayesian models is to consider the Fisher information matrix (FIM), which is defined as

$$M(\theta, \xi) = -\mathbb{E}_{p(y|\theta,\xi)} \left[\frac{\partial^2}{\partial \theta^2} \log p(y|\theta, \xi) \right] \tag{31}$$

where $\partial^2/\partial\theta^2$ denotes the Hessian when θ is a vector. The FIM has two important properties that motivate its use to extend the alphabetic criteria:

1. the FIM for the linear regression model is proportional to $(\xi^\top \xi)$;
2. the inverse FIM is related to the asymptotic covariance matrix of the Bayesian posterior by the Bernstein–von Mises Theorem (Van der Vaart, 2000).

For non-linear models, the FIM generally depends on θ as well as ξ, so forming a criterion for ξ involves an integral over $p(\theta)$. For instance, Chaloner and Verdinelli (1995) gives a Bayesian non-linear version of D-optimality as

$$U_{\text{Bayesian-}D}(\theta, \xi) = \log\det M(\theta, \xi)^{-1}; \tag{32}$$

substituting this utility in equation (9), leads to the optimality condition

$$\xi^* = \arg\max_{\xi} \mathbb{E}_{p(\theta)}[\log\det M(\theta, \xi)^{-1}]. \tag{33}$$

Using the FIM is not the only way to generalise the alphabetic criteria to non-linear models. Indeed, Ryan et al. (2016) takes issue with the classical FIM approach, suggesting that "to qualify as a 'fully Bayesian design', one must obtain the design by using a design criterion that is a functional of the posterior distribution". Whilst the EIG satisfies this requirement, the FIM extensions of the alphabetic criteria do not.

An approach to generalising the alphabetic criteria that is consistent with Ryan's definition of 'fully Bayesian' is to look at the covariance matrix of the Bayesian posterior $\text{Cov}_{p(\theta|y,\xi)}[\theta]$, which depends on ξ and y and is a functional of the posterior. For example, Ryan et al. (2016) mention two scalar objectives that can arise from this covariance matrix. One is termed the Bayesian D-posterior precision

$$U_{D\text{-precision}}(\xi, y) = \frac{1}{\det \text{Cov}_{p(\theta'|y,\xi)}[\theta']} \tag{34}$$

the other is quadratic loss

$$U_Q(\xi, y, \theta) = (\theta - \hat{\theta}(y, \xi))^\top A(\theta - \hat{\theta}(y, \xi)) \tag{35}$$

for some matrix A and for some posterior functional estimate $\hat{\theta}(y, \xi)$ of θ, such as the posterior mean. Both can be applied in the general framework of equation (9).

3.3 Other criteria

For a review of less common Bayesian experimental design criteria used within the statistics community, we recommend the reviews Chaloner and Verdinelli (1995) and Ryan et al. (2016).

4 Computational methods for one-step design

In this section, we focus on methods that have been utilised to solve the EIG maximisation problem

$$\begin{aligned} \xi^* &= \arg\max_{\xi\in\Xi} \mathcal{I}(\xi) \\ &= \arg\max_{\xi\in\Xi} \mathbb{E}_{p(y|\xi)}\left[\mathbb{E}_{p(\theta|\xi,y)}[\log p(\theta|\xi, y)] - \mathbb{E}_{p(\theta)}[\log p(\theta)]\right]. \end{aligned} \tag{36}$$

We further restrict ourselves to one-step experimental design, with sequential and adaptive design being left to Sec. 8, and to methods that have appeared broadly within the statistical community. Ryan et al. (2016) provides an excellent companion to this section.

Computational approaches to solving equation (36) can be broken down by

1. how they estimate the EIG objective $\mathcal{I}(\xi)$,
2. how they optimise ξ over the design space;

we tackle of each of these challenges in turn.

4.1 Point estimates of EIG

The EIG, $\mathcal{I}(\xi)$, is perhaps the most well studied criterion for Bayesian experimental design; it represents the expected reduction in Shannon entropy between the prior and posterior (see Sec. 3.1). The first step in utilising EIG for experimental design is to compute an estimate of the EIG for a single design ξ. Since $\mathcal{I}(\xi) = \mathbb{E}_{p(y|\xi)}\left[H[p(\theta)] - H[p(\theta|\xi, y)]\right]$ involves an expectation over $y \sim p(y|\xi)$ of the posterior entropy $H[p(\theta|y, \xi)]$, a direct approach to its estimation requires repeated computations of the posterior $p(\theta|y, \xi)$ with different simulated observations y. Given that calculating just one posterior can be intractable, it can readily be observed that EIG estimation is a computationally challenging problem. Foster et al. (2019) used the term 'doubly intractable' for computational problems of this kind.

A critical distinction when computing the EIG is whether the model has an explicit or an implicit likelihood (see Sec. 2.1 for a definition). In general, the explicit likelihood case contains strictly more information about the model, and so results in an easier, yet still doubly intractable, computational problem for the EIG. The implicit likelihood case is more challenging still, as the unknown likelihood typically has to be estimated in some way. We review computational methods for EIG estimation in both cases.

4.1.1 Explicit likelihood models

MCMC To begin with methods for EIG estimation with an explicit likelihood, a natural approach that is mostly suited to low θ dimension problems, is to estimate the posterior using Markov Chain Monte Carlo (MCMC) (Andrieu et al., 2003). Unfortunately, MCMC only produces samples of the target density. This is problematic for EIG estimation, which also requires access to the posterior density $p(\theta|\xi, y)$. To overcome this, Heinrich et al. (2020) used MCMC to sample the posterior, and a Gaussian Mixture Model (Hastie et al., 2009, Sec. 6.8) to perform density estimation of the posterior. MCMC has also been applied to estimate non-EIG criteria for Bayesian experimental design (Wakefield, 1994; Han and Chaloner, 2004).

Importance sampling Another family of methods for EIG estimation is based on importance sampling. These methods begin with the key observation that estimating the posterior density is not actually required for EIG estimation, because we can write

$$\log \frac{p(\theta|\xi, y)}{p(\theta)} = \log \frac{p(y|\theta, \xi)}{p(y|\xi)}. \tag{37}$$

The approach of Cook et al. (2008); Ryan et al. (2014) is to estimate $p(y|\xi)$ using Monte Carlo samples from the prior, leading to the estimator

$$\log p(y|\xi) \approx \log\left(\frac{1}{N}\sum_{n=1}^{N} p(y|\theta_n, \xi)\right) \text{ where } \theta_1, \ldots, \theta_N \overset{\text{i.i.d.}}{\sim} p(\theta), \tag{38}$$

the other component of the likelihood ratio $p(y|\theta, \xi)/p(y|\xi)$ is the known likelihood. Cook et al. (2008); Ryan et al. (2014) then estimate $\mathbb{E}_{p(\theta|\xi,y)}[\log p(y|\theta, \xi)]$ for some fixed y by using importance sampling. Specifically, given some fixed y and the set of samples $\theta_1, \ldots, \theta_n$ drawn independently of $p(\theta)$, they use the estimator

$$\mathbb{E}_{p(\theta|\xi,y)}[\log p(y|\theta, \xi)] \approx \frac{1}{N}\sum_{n=1}^{N} \frac{p(y|\theta_n, \xi)}{\frac{1}{N}\sum_{p=1}^{N} p(y|\theta_p, \xi)} \log p(y|\theta_n, \xi). \tag{39}$$

The final estimator of EIG is formed by combining this estimator with the estimator in equation (38) for $\log p(y|\xi)$, and then taking the Monte Carlo integral over $y \sim p(y|\xi)$, giving

$$\mathcal{I}(\xi) \approx \frac{1}{M}\sum_{m=1}^{M}\left[\frac{1}{N}\sum_{n=1}^{N}\frac{p(y_m|\theta_n,\xi)}{\frac{1}{N}\sum_{p=1}^{N}p(y_m|\theta_p,\xi)}\log p(y_m|\theta_n,\xi) - \log\left(\frac{1}{N}\sum_{n=1}^{N}p(y_m|\theta_n,\xi)\right)\right] \tag{40}$$

where $y_1 \ldots, y_m \overset{\text{i.i.d.}}{\sim} p(y|\xi)$ and $\theta_1, \ldots, \theta_n \overset{\text{i.i.d.}}{\sim} p(\theta)$ are independent.

Monte Carlo and Nested Monte Carlo Hamada et al. (2001); Ryan (2003) considered a closely related family of estimators. They also used equation (37) to avoid computing posterior densities. Unlike Cook et al. (2008); Ryan et al. (2014), they observed that $p(y|\xi)p(\theta|\xi,y) = p(\theta)p(y|\theta,\xi)$, allowing them to write the EIG as

$$\mathcal{I}(\xi) = \mathbb{E}_{p(\theta)p(y|\theta,\xi)}\left[\log\frac{p(y|\theta,\xi)}{p(y|\xi)}\right]. \tag{41}$$

The only unknown quantity in the integrand here is $p(y|\xi)$. Assuming some estimator $\hat{p}(y|\xi)$ for $p(y|\xi)$, we have the Monte Carlo estimator

$$\mathcal{I}(\xi) \approx \frac{1}{N}\sum_{n=1}^{N}\log\frac{p(y_n|\theta_n,\xi)}{\hat{p}(y_n|\xi)} \text{ where } \theta_n, y_n \overset{\text{i.i.d.}}{\sim} p(\theta)p(y|\theta,\xi). \tag{42}$$

In Hamada et al. (2001), $\hat{p}$ was computed by numerical integration, for a low dimensional θ. In Ryan (2003), two approaches for $\hat{p}$ were considered—the first was a Laplacian approximation using the posterior mode $\hat{\theta}$. The second was to use to an inner Monte Carlo estimation step to estimate $p(y|\xi)$ as in equation (38). This latter approach, also considered by Myung et al. (2013); Rainforth (2017), results in the double loop, or Nested Monte Carlo (NMC) estimator of EIG

$$\hat{\mathcal{I}}_{NMC}(\xi) = \frac{1}{N}\sum_{n=1}^{N}\log\frac{p(y_n|\theta_n,\xi)}{\frac{1}{M}\sum_{m=1}^{M}p(y_n|\theta'_m,\xi)} \text{ where } \theta_n, y_n \overset{\text{i.i.d.}}{\sim} p(\theta)p(y|\theta,\xi), \theta'_m \overset{\text{i.i.d.}}{\sim} p(\theta). \tag{43}$$

The asymptotic properties of this estimator were studied by Rainforth et al. (2018); Zheng et al. (2018); Beck et al. (2018), showing that $\hat{\mathcal{I}}_{NMC}(\xi)$ converges to $\mathcal{I}(\xi)$ with asymptotic error $\mathcal{O}(N^{-1}) + \mathcal{O}(M^{-2})$. Hence, it is optimal to set $M \propto \sqrt{N}$.

Laplace approximation Another important line of work (Lewi et al., 2009; Cavagnaro et al., 2010; Long et al., 2013) uses a Laplace approximation to the posterior to estimate the posterior entropy. The Laplace estimate uses the following Taylor expansion of a scalar function about a point $\hat{\theta}$

$$f(\theta) \approx f(\hat{\theta}) + (\theta-\hat{\theta})^\top \left.\frac{\partial f}{\partial\theta}\right|_{\hat{\theta}} + (\theta-\hat{\theta})^\top \left.\frac{\partial^2 f}{\partial\theta^2}\right|_{\hat{\theta}}(\theta-\hat{\theta}). \tag{44}$$

If we apply this approximation to the log posterior density $f(\theta) = \log p(\theta|\xi,y) = \log p(\theta) + \log p(y|\theta,\xi) + C$ at a point $\hat{\theta}$ for which the log posterior density has zero gradient, then we find the following Gaussian approximation

$$\log p(\theta|\xi,y) \approx (\theta-\hat{\theta})^\top\hat{\Sigma}^{-1}(\theta-\hat{\theta}) + C' \text{ where } \hat{\Sigma}^{-1} = \left.\frac{\partial^2\log\left(p(\theta)p(y|\theta,\xi)\right)}{\partial\theta^2}\right|_{\hat{\theta}}. \tag{45}$$

One advantage of this approach is that the entropy of this Gaussian approximation is known in closed form. A drawback is that the Laplace approximation makes a strong structural assumption about the posterior. This was partially relaxed by Long (2021), who considered a multi-modal Laplace approximation. Another approach is to combine Laplace estimation and importance sampling (Ryan et al., 2015). Finally, Beck et al. (2018) analysed the standard Laplace estimator, and further proposed combining Laplace importance sampling with the NMC estimator.

4.1.2 Implicit likelihood models

Approximating the likelihood In some models, the likelihood $p(y|\theta, \xi)$ can be computed, but it is too expensive to be used in extensive calculation. The approach of Huan and Marzouk (2013) is to approximate the likelihood using a polynomial chaos expansion. Here, it is necessary to use a small number of evaluations of the likelihood to compute the polynomial chaos coefficients, but once this is done, the surrogate polynomial chaos approximate likelihood can be used in place of the true likelihood for all other calculations. Huan and Marzouk (2013) specifically use the polynomial chaos approximation within a NMC estimator of the EIG.

Overstall and McGree (2020) also consider approximating the likelihood. They assume a parametric family for distributions over y with parameters ϕ, so that $y|\theta, \xi \sim \mathcal{H}_X(\phi_f(\theta, \xi))$. They estimate the function $\phi_f(\theta, \xi)$ using a Gaussian Process (Williams and Rasmussen, 2006), trained with data obtained by maximum likelihood estimation of ϕ_f. We note the close connections between this idea and Foster et al. (2019).

Approximate Bayesian Computation Approximate Bayesian Computation (ABC) (Csilléry et al., 2010) is a family of methods for performing inference without a tractable likelihood. In its simplest form, ABC simulates $(\tilde{\theta}_i, \tilde{y}_i)_{i=1}^N$ from the joint model $p(\theta, y|\xi)$. Given a metric ρ on $\mathcal{Y}$, a sample $\tilde{\theta}_i$ is accepted as a valid sample from $p(\theta|\xi, y)$ if

$$\rho(y, \tilde{y}_i) < \epsilon \tag{46}$$

for tolerance ϵ. Drovandi and Pettitt (2013), Hainy et al. (2016), Price et al. (2016) and Dehideniya et al. (2018) have applied ABC within the context of Bayesian experimental design.

LFIRE Another more recent approach to inference in intractable likelihood models is Likelihood-free Inference by Ratio Estimation (LFIRE) (Thomas et al., 2016). This method uses logistic regression to approximate the *likelihood ratio*

$$r(\xi, \theta, y) = \frac{p(y|\theta, \xi)}{p(y|\xi)}. \tag{47}$$

LFIRE was applied in a Bayesian experimental design context by Kleinegesse and Gutmann (2018) who sampled $(\theta_i, y_i)_{i=1}^N$ from the joint model and formed the Monte Carlo estimate of the EIG

$$\mathcal{I}(\xi) \approx \frac{1}{N} \sum_{i=1}^{N} \hat{r}(\xi, \theta_i, y_i). \tag{48}$$

4.2 Optimisation of EIG

We now turn to the problem of optimising the EIG over the design space Ξ.

4.2.1 Discrete design space

For a small, discrete design space, the simplest option is to form separate estimates of $\mathcal{I}(\xi)$ for each $\xi \in \Xi$, and choose the design with the highest estimated EIG. This approach was taken by Carlin et al. (1998); Palmer and Müller (1998) and others. Vincent and Rainforth (2017) dynamically allocated resources between different discrete designs using ideas from the theory of bandit optimisation (Neufeld et al., 2014). In essence, this approach provides more accurate EIG estimates for designs that are likely to be optimal, spending less time on designs that are not promising.

4.2.2 Continuous design space

Discretisation Perhaps the simplest approach to continuous design optimisation is to discretise the design space, for example using uniformly or log-uniformly spaced points (Ryan, 2003; van den Berg et al., 2003; Watson, 2017; Vincent and Rainforth, 2017). Alternatively, a discrete set of candidate designs can be chosen

by hand by the experimenter, and each evaluated (Han and Chaloner, 2004; Terejanu et al., 2012; Lyu et al., 2019).

Curve fitting Given a finite set of randomly sampled designs ξ_i with EIG estimates $\hat{\mathcal{I}}(\xi_i)$, Müller and Parmigiani (1995) proposed a curve fitting approach that fits a regression model to this data. The optimal design is then estimated as the optimum of the fitted regression model.

Bayesian optimisation Beyond simple curve fitting, Bayesian Optimisation (BO) (Snoek et al., 2012) is a well-established method for gradient-free optimisation. Like any other curve fitting approach, BO fits a model, specifically a Gaussian Process (GP), (Williams and Rasmussen, 2006) to the observed data $(\xi_i, \hat{\mathcal{I}}(\xi_i))$. However, BO iteratively suggests new designs at which to estimate the EIG, in order to efficiently seek the optimal design. We fully discuss BO and its connection with Bayesian experimental design itself in Sec. 7. For the purposes of solving the EIG optimisation problem, equation (36), we treat BO as a black box optimisation algorithm. The application of BO to optimising EIG over the design space was explored by Kleinegesse and Gutmann (2018); Foster et al. (2019); von Kügelgen et al. (2019).

Co-ordinate exchange The classical co-ordinate exchange algorithm for optimising design was proposed by Meyer and Nachtsheim (1995). Overstall and Woods (2017) proposed Approximate Co-ordinate Exchange. This is a two phase optimisation algorithm specifically designed for Bayesian experimental design. In the first phase, designs are optimised co-ordinate-wise by fitting a one-dimensional GP to the EIG surface for each co-ordinate in turn, with other elements of the design held fixed, and selecting the optimal value for that co-ordinate. In the second phase, different co-ordinates of the design are aggregated using a point exchange algorithm (Meyer and Nachtsheim, 1995; Atkinson et al., 2007).

Optimisation by sampling Clyde et al. (1996) proposed an approach to optimising the design that uses algorithms for sampling unnormalised densities. Their approach applies to any utility $U(\theta, \xi, y) > 0$ in the framework of equation (9) that does not use an extra decision variable δ. The authors define an augmented probability model on $\Xi \times \Theta \times \mathcal{Y}$ by

$$h(\xi, \theta, y) \propto p(\theta)p(y|\theta, \xi)U(\theta, \xi, y). \tag{49}$$

The marginal distribution for ξ is then

$$h(\xi) \propto \mathbb{E}_{p(\theta)p(y|\theta,\xi)}[U(\theta, \xi, y)], \tag{50}$$

this guarantees that high probability regions for ξ correspond to regions with a large utility. The core approach, then, is to sample from the joint density $h(\xi, \theta, y)$ using a technique such as MCMC—Clyde et al. (1996) used the Metropolis–Hastings algorithm (Hastings, 1970). MCMC on $h(\theta, \xi, y)$ was also used by Bielza et al. (1999); Müller (2005). Cook et al. (2008); Drovandi and Pettitt (2013) used the MCMC technique, and fitted a density estimator to the MCMC samples to improve their estimation of the optimal design. Ryan et al. (2014) applied MCMC in combination with dimensionality reduction on the latent space to avoid problems with MCMC in higher dimensions.

An extension of this idea, inspired by simulated annealing (Van Laarhoven and Aarts, 1987), is to include the utility contributions from J independent (θ, y) pairs, to create an unnormalised density on $\Xi \times \Theta^J \times \mathcal{Y}^J$

$$h_J(\xi, \theta_{1:J}, y_{1:J}) = \prod_{j=1}^{J} p(\theta_j)p(y_j|\theta_j, \xi)U(\theta_j, \xi, y_j). \tag{51}$$

One can see that for larger J, the probability mass concentrates more strongly around the optimal ξ. The simulated annealing mechanism is applied by *increasing J during the course of optimisation.* This approach has been applied by Müller et al. (2004); Müller (2005); Stroud et al. (2001); Cook et al. (2008).

Alternatively, one can sample $h_J(\xi, \theta_{1:J}, y_{1:J})$ using Sequential Monte Carlo (SMC) (Doucet et al., 2000). This was the approach taken by Amzal et al. (2006); Kuck et al. (2006).

Evolutionary algorithms Another approach to solving equation (36) is to optimise over the design space using evolutionary algorithms (Eiben et al., 2003). Hamada et al. (2001) applied genetic algorithms to this problem, and Price et al. (2018) proposed the Induced Natural Selection Heuristic (INSH) method to optimise the design.

Gradient-based optimisation Huan and Marzouk (2014) considered gradient-based methods for solving the EIG optimisation problem. They considered the Robbins–Monroe stochastic gradient descent (SGD) (Robbins and Monro, 1951) algorithm applied to the NMC estimator of the EIG, equation (43), resampling θ_n, y_n and θ'_m at each iteration. They also considered applying the SAA-BFGS algorithm (Fletcher, 2013) to the NMC estimator, without resampling at each iteration. Carlon et al. (2020) considered gradient optimisation of both NMC and Laplace estimators of the EIG using SGD.

Other methods Huan and Marzouk (2013) proposed the Nelder–Mead simplex method (Nelder and Mead, 1965), a gradient-free optimisation algorithm, and simultaneous perturbation stochastic approximation (Spall, 1998) as two alternative optimisation algorithms for Bayesian experimental designs.

5 Bayesian Active Learning

Active learning allows a learning algorithm to "choose the data from which it learns" (Settles, 2009). In the Bayesian setting, the learning algorithm is a Bayesian model. In its most abstract form, then, Bayesian Active Learning is identical to Bayesian Experimental Design, but with different vocabulary: designs ξ are referred to as queries, observations y are referred to as labels and are often provided by a human labeller, the design criterion is referred to as the acquisition function. Queries are selected to maximise the acquisition function, typically in an iterative process.

Pool-based active learning However, this abstract similarity disguises the common differences in applications of active learning and experimental design. One hugely important sub-field of active learning, including Bayesian active learning, is *pool-based active learning* (Lewis and Gale, 1994). Here, the design space Ξ consists of unlabelled examples (such as images or sentences), the observation y is a human-provided label that corresponds to the unlabelled instance ξ, and the model is a classifier with parameters θ that predicts y from ξ. Pool-based active learning also applies less commonly to regression problems, for which y is a continuous label.

Sequential active learning with greedy acquisition In Sections 3 and 4, we focused on one-step design in which we begin with a prior $p(\theta)$, select a design ξ, obtain outcome y, and the experiment terminates. In active learning, we rarely want to acquire just one label or one batch of labels—the true power of the framework is apparent in a sequential setting (Lewis and Gale, 1994). This means that we pick design ξ_1 obtaining label y_1, then choose ξ_2 and receive label y_2, and so on. The dataset that we have after t experiments is $\mathcal{D}_t = \{(\xi_1, y_1), \dots, (\xi_t, y_t)\}$. A simple approach to the sequential problem that is adopted in almost all of active learning (Gal et al., 2017) is *greedy acquisition*. In short, this strategy picks the next design to maximise the utility of the next label, without any consideration of how this will affect future queries.

However, it is still essential to incorporate all existing data $\mathcal{D}_t$ into the model before making this choice. To do this, we use the posterior[3] given existing data $p(\theta|\mathcal{D}_t)$ in place of the original prior $p(\theta)$. For the EIG, for example, at each step we would choose the design that maximises

$$\mathcal{I}(\xi; \mathcal{D}_t) = \mathbb{E}_{p(y|\xi,\mathcal{D}_t)}\left[\mathbb{E}_{p(\theta|\xi,y,\mathcal{D}_t)}[\log p(\theta|\xi, y, \mathcal{D}_t)] - \mathbb{E}_{p(\theta|\mathcal{D}_t)}[\log p(\theta|\mathcal{D}_t)]\right] \tag{52}$$

[3]In active learning, we make the assumption that θ represents the full set of model parameters (see Sec. 6).

Algorithm 1 Pool-based Bayesian active learning with greedy acquisition

Require: Acquisition function α, prior $p(\theta)$ on model weights, pool Ξ, initial dataset $\mathcal{D}_0$ may be empty.
for step $t = 1, \ldots, T$ **do**
Find $\xi_t = \arg\max_{\xi \in \Xi} \alpha(\xi; \mathcal{D}_{t-1})$ by scoring each unlabelled element of the pool
Obtain label y_t for query ξ_t
Set $\mathcal{D}_t = \mathcal{D}_{t-1} \cup \{(\xi_t, y_t)\}$ and retrain model to compute $p(\theta|\mathcal{D}_t)$
end for

where $p(y|\xi, \mathcal{D}_t) = \mathbb{E}_{p(\theta|\mathcal{D}_t)}[p(y|\theta, \xi)]$. The high-level framework of greedy, sequential pool-based Bayesian active learning with a general acquisition function α is summarised in Algorithm 1. We discuss the theory of sequential experimentation in more detail in Sec. 8.

5.1 Acquisition functions

5.1.1 Bayesian Active Learning by Disagreement

A key point of intersection between Bayesian active learning and Bayesian experimental design is the Bayesian Active Learning by Disagreement (BALD) score (Houlsby et al., 2011), a widely adopted acquisition function within Bayesian Active Learning.

Proposition 9 (Houlsby et al. (2011))**.** *The BALD score is equivalent to the EIG.*

Proof. The BALD score is the mutual information between θ and y, but typically rearranged as

$$\alpha_{\text{BALD}}(\xi; \mathcal{D}_t) = \mathbb{E}_{p(\theta|\mathcal{D}_t)}\left[H[p(y|\xi, \mathcal{D}_t)] - H[p(y|\xi, \theta, \mathcal{D}_t)]\right]. \tag{53}$$

We have

$$= \mathbb{E}_{p(\theta|\mathcal{D}_t)}\left[-\mathbb{E}_{p(y|\xi,\mathcal{D}_t)}[\log p(y|\xi, \mathcal{D}_t)] + \mathbb{E}_{p(y|\xi,\theta,\mathcal{D}_t)}[\log p(y|\xi, \theta, \mathcal{D}_t)]\right] \tag{54}$$

$$= \mathbb{E}_{p(\theta|\mathcal{D}_t)p(y|\xi,\theta,\mathcal{D}_t)}\left[-\log p(y|\xi, \mathcal{D}_t) + \log p(y|\xi, \theta, \mathcal{D}_t)\right] \tag{55}$$

$$= \mathbb{E}_{p(\theta|\mathcal{D}_t)p(y|\xi,\theta,\mathcal{D}_t)}\left[\log \frac{p(y|\xi, \theta, \mathcal{D}_t)}{p(y|\xi, \mathcal{D}_t)}\right] \tag{56}$$

applying Bayes Theorem gives

$$= \mathbb{E}_{p(\theta|\mathcal{D}_t)p(y|\xi,\theta,\mathcal{D}_t)}\left[\log \frac{p(\theta|\xi, y, \mathcal{D}_t)}{p(\theta|\mathcal{D}_t)}\right] \tag{57}$$

$$= \mathbb{E}_{p(y|\xi,\mathcal{D}_t)}\left[\mathbb{E}_{p(\theta|\xi,y,\mathcal{D}_t)}[\log p(\theta|\xi, y, \mathcal{D}_t)] - \mathbb{E}_{p(\theta|\mathcal{D}_t)}[\log p(\theta|\mathcal{D}_t)]\right] = \mathcal{I}(\xi; \mathcal{D}_t). \tag{58}$$

Note this is essentially the same proof as Proposition 6. □

The BALD score can be utilised directly in Algorithm 1. One important feature of writing EIG in BALD form is that it only depends on the actual experimental observation y, and does not require a probability density on θ. This can be important if we do not have a closed form density for θ either in the prior $p(\theta)$ or in the posterior $p(\theta|\mathcal{D}_t)$. This is particularly useful in active learning, where we may consider particularly complex models with high-dimensional θ.

In Deep Bayesian Active Learning (Gal et al., 2017), for instance, the model that predicts y from ξ is a neural network with parameters θ. In order to treat this model in a Bayesian manner, methods for Bayesian deep learning must be utilised. Gal et al. (2017) specifically used Dropout as a way of estimating prior and posterior distributions on θ (Gal and Ghahramani, 2016). Here, fitting $p(\theta|\mathcal{D}_t)$ amounts to retraining the network with Dropout. Beluch et al. (2018) and Pop and Fulop (2018) used a simple ensemble of models, treating different members of the ensemble as posterior samples of θ. To fit $p(\theta|\mathcal{D}_t)$, each deterministic model in the ensemble is retrained separately.

A key computational insight when estimating $\mathcal{I}(\xi)$ for a classification model in which the observation space $\mathcal{Y}$ is finite was made by Houlsby et al. (2011); Gal et al. (2017). We have

$$\mathcal{I}(\xi) = \sum_{y \in \mathcal{Y}} \mathbb{E}_{p(\theta)} \left[p(y|\theta, \xi) \log \frac{p(y|\theta, \xi)}{\mathbb{E}_{p(\theta)}[p(y|\theta, \xi)]} \right] \tag{59}$$

which can simply be estimated with Monte Carlo using samples $\theta_1, \ldots, \theta_N \sim p(\theta)$. The same idea applied when we have $p(\theta|\mathcal{D}_t)$ in place of $p(\theta)$. This estimator was also used by Vincent and Rainforth (2017), who observed that, unlike the NMC estimator of equation (43), this estimator converges at the standard Monte Carlo rate with error $\mathcal{O}(N^{-1/2})$. This speed-up is a consequence of being able to sum over $\mathcal{Y}$.

BatchBALD In the pool-based active learning setting with a discrete pool of size p, each acquisition involves computing the BALD score for every element of the pool and choosing the best one (Algorithm 1), which is an $\mathcal{O}(p)$ operation. Kirsch et al. (2019) considered the problem of batch active learning, in which designs are k-subsets of the pool. This means that, at each iteration of active learning, k different unlabelled examples will be selected and labelled. Naively scoring each k-subset of the unlabelled pool costs $\binom{p}{k}$, which rapidly becomes prohibitive. BatchBALD instead creates the design by greedily adding elements from the pool one at a time, giving a more efficiently scalable algorithm. This approach can be justified theoretically using the notion of submodularity—see Sec. 8.1.1.

5.1.2 Other acquisition functions

Within the Bayesian active learning framework, a range of other acquisition functions and computational methods have been proposed. It is possible to extend most common *non-Bayesian* acquisition functions for use with Bayesian models. These non-Bayesian acquisition rules are generally a function of the predictive distribution $p(y|\xi, \mathcal{D}_t)$. When using a Bayesian model we can use the Bayesian marginal (posterior predictive) $p(y|\xi, \mathcal{D}_t) = \mathbb{E}_{p(\theta|\mathcal{D}_t)}[p(y|\theta, \xi)]$ in place of the deterministic predictive distribution that arises in non-Bayesian models. Standard acquisition functions such as uncertainty sampling (Lewis and Gale, 1994), margin sampling (Scheffer et al., 2001), and variation ratios (Freeman, 1965) can be therefore be employed in this context. Of particular note is the maximum entropy sampling method (Shannon, 1948; Settles and Craven, 2008), which uses

$$\alpha_{\text{Entropy}}(\xi; \mathcal{D}_t) = H[p(y|\xi, \mathcal{D}_t)]. \tag{60}$$

As shown in Proposition 6, this approach is equivalent to EIG maximisation when the entropy $H[p(y|\theta, \xi)]$ does not depend on ξ. This can be interpreted as saying that, given the correct model, the level of noise is uniform across all examples in the pool Ξ. For instance, we could assume that every example has a true label that a human will assign with 100% accuracy. However, maximum entropy sampling (and, in general, rules based on uncertainty in the predictive distribution $p(y|\xi, \mathcal{D}_t)$) break down when there are designs ξ which are very ambiguous, e.g. the correct label is missing from the taxonomy. Maximum entropy and related acquisition rules can become fixated on ambiguous queries.

Active learning has also considered Bayesian-specific acquisition functions. Kendall et al. (2015) proposed the mean standard deviation (Mean STD) acquisition rule for classification models. Define $\sigma_y(\xi; \mathcal{D}_t)$ as the standard deviation over $\theta|\mathcal{D}_t$ of the probability of example ξ being assigned to class y, i.e.

$$\sigma_y(\xi; \mathcal{D}_t) = \sqrt{\text{Var}_{p(\theta|\mathcal{D}_t)}[p(y|\theta, \xi)]}, \tag{61}$$

then the MeanSTD acquisition function is,

$$\alpha_{\text{MeanSTD}}(\xi; \mathcal{D}_t) = \frac{1}{|\mathcal{Y}|} \sum_{y \in \mathcal{Y}} \sigma_y(\xi; \mathcal{D}_t). \tag{62}$$

Following Bayesian decision theory, Roy and McCallum (2001) considered minimising the Bayes posterior risk, focusing on log loss and 0/1 loss. Kapoor et al. (2007) considered a range of Bayesian acquisition functions for binary classification, focusing on a score which combines the mean and variance of the prediction. Yang et al. (2012) applied Bayesian active learning to metric learning, and used an acquisition function based on maximum entropy.

6 Embedded models

So far, we have assumed that θ, the parameters of interest, and θ, the full set of model parameters, are one and the same. In this section, we explore the case in which the parameters of interest and the full set of model parameters are different. Model selection (Vanlier et al., 2014; Drovandi et al., 2014), in which we are only interesting in deciding which model is correct and not interested in learning the exact model parameters, is one important example of this setting. Bayesian optimisation (Sec. 7) is also an example in which we have a probability model for an unknown function, but we are only interest in learning the location of the maximum of that function.

For this more general case, we assume that the model is fully specified by a set of parameters ψ, and that our parameters of interest θ are a function of the full parameter set $\theta = f_\theta(\psi)$. In this case, the full joint distribution of the model is $p(\psi, y|\xi)$, and we obtain a joint over $\Theta \times \mathcal{Y}$ by integrating

$$p(\theta, y|\xi) = \int_{f_\theta^{-1}(\{\theta\})} p(\psi, y|\xi)\, d\psi. \tag{63}$$

Semi-implicit likelihood The embedded model setting allows us to extend our discussion of explicit and implicit models (Sec. 2.1). It could be the case that we do have an explicit prior for ψ and an explicit likelihood $p(y|\psi, \xi)$ for the observation y given the full set of parameters ψ. Then the likelihood $p(y|\theta, \xi)$ is given by

$$p(y|\theta, \xi) = \int_{f_\theta^{-1}(\{\theta\})} p(y|\psi, \xi)\, d\psi, \tag{64}$$

and the prior is given by

$$p(\theta) = \int_{f_\theta^{-1}(\{\theta\})} p(\psi)\, d\psi. \tag{65}$$

First, note that $p(\theta, y|\xi) \neq p(\theta)p(y|\theta, \xi)$ in an embedded model. Second, computing one or both of these integrals may be an intractable computation. We use the term *semi-implicit* for this case in which the likelihood or prior for ψ is explicit, but the likelihood or prior for θ involves an intractable integral. So a semi-implicit likelihood is one which is formed as an integral of an explicit likelihood $p(y|\psi, \xi)$ and a semi-implicit prior is one which is an integral of explicit prior $p(\psi)$.

Exchangeability In an exchangeable embedded model, it is no longer true that different experiments are independent *conditional on* θ. Intuitively, the reason for this is that one experiment gives us information about *all of* ψ. Without extra assumptions, information from the first experiment tells us something about ψ even when we condition on θ, and this influences the predictive distribution for the second experiment. More formally, the factorisation equation (3) must be replaced by a factorisation conditional on ψ, and the natural assumption to make is that experiments are independent conditional on ψ

$$p(\psi, y_{1:T}|\xi_{1:T}) = p(\psi) \prod_{t=1}^{T} p(y_t|\psi, \xi_t). \tag{66}$$

Sequential learning with greedy acquisition One of the consequences of equation (66) is that Algorithm 1 is not quite correct for an embedded model. Specifically, between iterations, it is necessary to update the full model on ψ by fitting $p(\psi|\mathcal{D}_t)$, it is not enough to update beliefs about θ.

6.1 Expected Information Gain for embedded models

The EIG can naturally extend to the case of embedded models. The definition of information gain on the parameter of interest θ remains the same: $U_{\mathcal{I}}(\xi, y) = \mathbb{E}_{p(\theta|\xi,y)}[\log p(\theta|y, \xi)] - \mathbb{E}_{p(\theta)}[\log p(\theta)]$. When we take

the expectation over y, however, we use the Bayesian marginal that integrates over all of ψ, i.e. $p(y|\xi) = \mathbb{E}_{p(\psi)}[p(y|\psi,\xi)]$, to give

$$I(\xi) = \mathbb{E}_{p(\psi)p(\theta|\psi)p(y|\psi,\xi)}\left[\log \frac{p(\theta|y,\xi)}{p(\theta)}\right]. \tag{67}$$

where $p(\theta|\psi)$ is a delta function on $f_\theta(\psi)$. This is different to the definition in equation (17) because the expectation is taken over $p(\theta, y|\xi) \neq p(\theta)p(y|\theta,\xi)$ for an embedded model. The EIG in an embedded model can also be expressed in BALD form (Proposition 9) as

$$\begin{aligned} I(\xi) &= H[p(y|\xi)] - \mathbb{E}_{p(\theta)}[H[p(y|\theta,\xi)]] &(68)\\ &= H\left[\mathbb{E}_{p(\psi)}[p(y|\psi,\xi)]\right] - \mathbb{E}_{p(\psi)p(\theta|\psi)}[H[p(y|\theta,\xi)]], &(69)\end{aligned}$$

where the second line emphasises the difference with the standard case.

6.2 Computational methods for semi-implicit likelihood models

The NMC estimator of the EIG (Ryan, 2003) can be extended to the semi-implicit case. The central idea is that we form a Monte Carlo estimator of both $p(y|\theta,\xi)$ and $p(y|\xi)$ using appropriate Monte Carlo integrals over ψ. As in the standard NMC estimator, we have

$$p(y|\xi) = \mathbb{E}_{p(\psi)}[p(y|\psi,\xi)] \approx \frac{1}{M}\sum_{m=1}^{M} p(y|\psi_m,\xi) \text{ where } \psi_1,\ldots,\psi_M \overset{\text{i.i.d.}}{\sim} p(\psi). \tag{70}$$

For $p(y|\theta,\xi)$, we need access to samples from the distribution $p(\psi|\theta)$. Then,

$$p(y|\theta,\xi) = \mathbb{E}_{p(\psi|\theta)}[p(y|\psi,\xi)] \approx \frac{1}{M}\sum_{m=1}^{M} p(y|\psi_m,\xi) \text{ where } \psi_1,\ldots,\psi_M \overset{\text{i.i.d.}}{\sim} p(\psi|\theta). \tag{71}$$

Combining, we have the semi-implicit NMC estimator of EIG

$$\hat{I}_{\text{SI-NMC}}(\xi) = \frac{1}{N}\sum_{n=1}^{N}\left[\log\left(\frac{1}{M}\sum_{m=1}^{M} p(y_n|\psi_{nm},\xi)\right) - \log\left(\frac{1}{M}\sum_{m=1}^{M} p(y_n|\psi_m,\xi)\right)\right] \tag{72}$$

where $\theta_n, y_n \overset{\text{i.i.d.}}{\sim} p(\theta, y|\xi)$, $\psi_m \overset{\text{i.i.d.}}{\sim} p(\psi)$ and $\psi_{nm} \overset{\text{i.i.d.}}{\sim} p(\psi|\theta_n)$.

Ma et al. (2018) considered information acquisition for imputation in a semi-implicit setting. They used a Partial VAE which facilitated estimation of the EIG using the conditional independence assumptions of the model. Extending this, Gong et al. (2019) considered a similar active imputation scenario. They used $\hat{I}_{\text{SI-NMC}}(\xi)$ to estimate an information criterion for experimental design. In their probabilistic model, they had $\psi = (\theta, z)$ with $p(\psi) = p(\theta)p(z)$. When conditioning on data $\mathcal{D}_t$, they used approximate inference in which the independence of θ and z was maintained. Under these conditions, sampling $p(\psi|\theta)$ amounted to fixing θ and taking new, independent samples of z. A simplified form of their estimator is

$$\hat{I}_{\text{Icebreaker}}(\xi) = \frac{1}{N}\sum_{n=1}^{N}\left[\log\left(\frac{1}{M}\sum_{m=1}^{M} p(y_n|\theta_n,z_m,\xi)\right) - \log\left(\frac{1}{ML}\sum_{m=1}^{M}\sum_{\ell=1}^{L} p(y_n|\theta_\ell,z_m,\xi)\right)\right] \tag{73}$$

where $\theta_n, y_n \overset{\text{i.i.d.}}{\sim} p(\theta, y|\xi)$, $z_m \overset{\text{i.i.d.}}{\sim} p(z)$ and $\theta'_\ell \overset{\text{i.i.d.}}{\sim} p(\theta)$. Overstall and Woods (2017) considered almost the same setting when estimating the EIG utility. Specifically, they considered a semi-implicit case in which ψ can be partitioned into parameters of interest and *independent* nuisance parameters, and used this semi-implicit NMC estimator for the EIG.

7 Bayesian Optimisation

Bayesian optimisation (BO) (Snoek et al., 2012; Shahriari et al., 2015) considers the problem of finding the maximiser of an unknown objective function

$$\xi^* = \arg\max_{\xi \in \Xi} f(\xi). \tag{74}$$

To deal with the unknown function in a Bayesian manner, we consider a statistical model for f with prior $p(f)$. We assume that we can obtain relatively expensive measurements from the true function f at design points ξ. These measurements may be corrupted by noise, meaning that we obtain observations

$$y|\xi, f \sim p(y|f(\xi)), \tag{75}$$

for example, $y = f(\xi) + \varepsilon$ for $\varepsilon \sim N(0, \sigma^2)$.

BO can naturally be cast within the framework of Bayesian experimental design. We have designs ξ and observations y connected by the Bayesian model on f and the noise model. The missing piece is to specify the parameter of interest θ. The parameter of interest is not f, because Bayesian optimisation is explicitly concerned with *maximising* f, meaning that any information about f in regions where it is well below its maximum is not useful. The most common formulation is to take to be the location of the maximiser of f (Hernández-Lobato et al., 2014), i.e. $\theta = \arg\max_{\xi \in \Xi} f(\xi)$. The fact that θ is not all of f means that BO is not an explicit likelihood (Sec. 2.1) experimental design problem, nor does it fit into the framework of Bayesian active learning (Sec. 5). BO is experimental design for an embedded model (Sec. 6), with the function f playing the role of the richer parameter set ψ. We will see that BO has its own character with a wide range of algorithms that apply specifically to the optimisation problem.

To set up a BO system, we begin by specifying a Bayesian model for f with prior $p(f)$, and a measurement noise model $p(y|f(\xi))$. We then specify an acquisition function that guides our choice of designs at which we should take measurements. The acquisition function in BO plays the same role as the design criterion in Bayesian experimental design and the acquisition function in active learning—we select the design that maximises the acquisition function to obtain new measurements of f. As in active learning, BO typically adopts the *greedy acquisition* approach that was outlined in Sec. 5. The entire approach is summarised in Algorithm 2.

We begin by discussing common choices for the Bayesian model and acquisition function in BO. We focus specifically on the Entropy Search family of acquisition rules, highlighting the connection to experimental design with EIG.

7.1 Bayesian models for optimisation

7.1.1 Parametric models

When the design space Ξ is discrete, the function f can be characterised by a finite number of latent variables. This case is closely connected to theory of multi-armed bandits (Lai and Robbins, 1985): we can view each $\xi \in \Xi$ as an 'arm' of a bandit in a casino. Each arm has an unknown payout, and the aim is to identify the best arm. Our mathematical set-up specifically relates the pure exploration scenario (Bubeck et al., 2009), in which final knowledge of the location of the best arm is important, but function evaluations during the course of Algorithm 2 are not. Finite-dimensional models such as the Beta-Bernoulli (Shahriari et al., 2015) and the Gaussian (Hoffman et al., 2014) have been applied in the bandit context.

Both Bayesian linear and generalised linear models have been utilised within the Bayesian optimisation context (Russo and Van Roy, 2014; Shahriari et al., 2015). In the bandit context, these models are applied by associating each bandit arm with a feature vector $\mathbf{x}_\xi$, and assuming that the arm payout depends on this feature vector. For a linear model, for example, we would assume $f(\xi) = \langle \mathbf{x}_\xi, \mathbf{w} \rangle$. These models can also be applied to optimisation over continuous design spaces. Snoek et al. (2015) considered Bayesian optimisation using a Bayesian neural network as the model for f; they specifically took an 'adaptive basis regression' approach that is only Bayesian on the last layer of the network.

Algorithm 2 Bayesian Optimisation (Shahriari et al., 2015)

Require: Acquisition function α, prior $p(f)$ on function, design space Ξ, initial dataset $\mathcal{D}_0$ may be empty.
for step $t = 1, \ldots, T$ **do**
Find $\xi_t = \arg\max_{\xi \in \Xi} \alpha(\xi; \mathcal{D}_{t-1})$
Obtain noisy measurement $y_t \sim p(y|f(\xi_t))$ at design ξ_t
Set $\mathcal{D}_t = \mathcal{D}_{t-1} \cup \{(\xi_t, y_t)\}$ and retrain the model to compute $p(f|\mathcal{D}_t)$
end for
Use $p(f|\mathcal{D}_T)$ to estimate the maximiser of f.

7.1.2 Nonparametric models

For continuous Bayesian optimisation, the Gaussian Process (GP) (Williams and Rasmussen, 2006) has proved an extremely popular Bayesian nonparametric model for the unknown function f (Osborne et al., 2009). The Gaussian process with a positive definite kernel k and mean function μ assumes the following multivariate Gaussian distribution for the finite-dimensional marginal distributions (Øksendal, 2003) of f

$$\begin{pmatrix} f(\xi_1) \\ \vdots \\ f(\xi_n) \end{pmatrix} \sim N\left(\begin{pmatrix} \mu(\xi_1) \\ \vdots \\ \mu(\xi_n) \end{pmatrix}, \begin{pmatrix} k(\xi_1, \xi_1) & \ldots & k(\xi_1, \xi_n) \\ \vdots & & \vdots \\ k(\xi_n, \xi_1) & \ldots & k(\xi_n, \xi_n) \end{pmatrix} \right). \tag{76}$$

Given a dataset of observations $\mathcal{D}_t = \{(\xi_i, y_i)\}_{i=1}^t$, the resulting posterior on f is also a Gaussian process. The mean and covariance structure of the posterior can be derived by computing the conditional form of equation (76), however, the necessary matrix computations come at cubic cost $\mathcal{O}(t^3)$; we refer to Williams and Rasmussen (2006) for full details. As a mark of its popularity, BO with a GP model for f has been implemented in several software frameworks, such as BoTorch (Balandat et al., 2020).

Within Bayesian optimisation, several extensions of the GP have also been considered as models for f. Different variants of *sparse* GPs have been proposed (Quinonero-Candela and Rasmussen, 2005; Snelson and Ghahramani, 2006; Lázaro-Gredilla et al., 2010), aiming to reduce the computational burden of using the standard GP conditioning formula. Calandra et al. (2016) combined GPs with feature learning to propose the Manifold GP.

Beyond the GP family, Hutter et al. (2013) also considered a random forest model for f, but found GPs to be preferable. Focusing on the application of hyperparameter optimisation, Bergstra et al. (2011) proposed the Tree-structured Parzen Estimator model for f that combines a tree-structured hierarchy with mixture modelling. Finally, Neiswanger et al. (2019) considered Bayesian optimisation in which an arbitrary probabilistic program is used as the model for f.

7.2 Acquisition functions

7.2.1 The Entropy Search family

To define an information-theoretic acquisition function for Bayesian optimisation, we want to gain information about the random variable $\theta = \arg\max_{\xi \in \Xi} f(\xi)$. To this end, Villemonteix et al. (2009) proposed Stepwise Uncertainty Reduction (SUR). This method aims to reduce posterior entropy in θ using the acquisition rule

$$\alpha_{\mathrm{SUR}}(\xi; \mathcal{D}_t) := -\mathbb{E}_{p(y|\xi, \mathcal{D}_t)}[H[p(\theta|\mathcal{D}_t \cup \{(\xi, y)\})]]. \tag{77}$$

where $p(y|\xi, \mathcal{D}_t) = \mathbb{E}_{p(f|\mathcal{D}_t)}[p(y|f(\xi))]$. In practice, Villemonteix et al. (2009) estimated the acquisition function by discretising θ and using a GP model for f. Hennig and Schuler (2012) considered a closely related acquisition function called Entropy Search (ES) that maximises the KL-divergence between the posterior on θ and a base measure $b(\theta)$. This gives the acquisition function

$$\alpha_{\mathrm{ES}}(\xi; \mathcal{D}_t) := \mathbb{E}_{p(y|\xi, \mathcal{D}_t)}[\mathrm{KL}[p(\theta|\mathcal{D}_t \cup \{(\xi, y)\}) \| b(\theta)]]. \tag{78}$$

The following Proposition, due to MacKay (1992), shows that these in information measures are equivalent, and are equivalent to the EIG.

Proposition 10 (MacKay (1992))**.** *Consider the general experimental design set-up of Sec. 3. The following acquisition functions all give the same optimal design*

$$\mathcal{I}(\xi) = \mathbb{E}_{p(y|\xi)}\left[\mathbb{E}_{p(\theta|\xi,y)}[\log p(\theta|\xi,y)] - \mathbb{E}_{p(\theta)}[\log p(\theta)]\right], \tag{79}$$
$$\mathcal{I}_2(\xi) = -\mathbb{E}_{p(y|\xi)}[H[p(\theta|\xi,y)]], \tag{80}$$
$$\mathcal{I}_3(\xi) = \mathbb{E}_{p(y|\xi)}[\mathrm{KL}[p(\theta|\xi,y)\|b(\theta)]] \tag{81}$$

where $p(y|\xi) = \mathbb{E}_{p(f)}[p(y|f(\xi))]$.

Proof. We have

$$\mathcal{I}(\xi) = \mathbb{E}_{p(y|\xi)}\left[\mathbb{E}_{p(\theta|\xi,y)}[\log p(\theta|\xi,y)] - \mathbb{E}_{p(\theta)}[\log p(\theta)]\right] \tag{82}$$
$$= \mathbb{E}_{p(y|\xi)}[-H[p(\theta|\xi,y)] + H[p(\theta)]] \tag{83}$$
$$= \mathcal{I}_2(\xi) + H[p(\theta)]. \tag{84}$$

and

$$\mathcal{I}(\xi) = \mathbb{E}_{p(y|\xi)}\left[\mathbb{E}_{p(\theta|\xi,y)}[\log p(\theta|\xi,y)] - \mathbb{E}_{p(\theta)}[\log p(\theta)]\right] \tag{85}$$
$$= \mathbb{E}_{p(y|\xi)}\left[\mathbb{E}_{p(\theta|\xi,y)}\left[\log \frac{p(\theta|\xi,y)}{b(\theta)}\right] - \mathbb{E}_{p(\theta)}\left[\log \frac{p(\theta)}{b(\theta)}\right]\right] \tag{86}$$
$$= \mathcal{I}_3(\xi) - \mathrm{KL}[p(\theta)\|b(\theta)]. \tag{87}$$

Since $H[p(\theta)]$ and $\mathrm{KL}[p(\theta)\|b(\theta)]$ do not depend on ξ, choosing ξ to maximise EIG is equivalent to maximising $\mathcal{I}_2$ and $\mathcal{I}_3$. □

Corollary 11. *Stepwise Uncertainty Reduction (Villemonteix et al., 2009) and Entropy Search (Hennig and Schuler, 2012) are equivalent to EIG maximisation when* $\theta = \arg\max_{\xi\in\Xi} f(\xi)$.

Proof. Note that $\mathcal{I}_2 = \alpha_{\mathrm{SUR}}$ and $\mathcal{I}_3 = \alpha_{\mathrm{ES}}$ when we replace the prior $p(\theta)$ with the posterior $p(\theta|\mathcal{D}_t)$. Since the result holds for a general experimental design set-up, it specifically holds in the BO case when $\theta = \arg\max_{\xi\in\Xi} f(\xi)$. □

Hernández-Lobato et al. (2014) proposed Predictive Entropy Search (PES). Like previous methods, PES uses the EIG (accounting for the embedded model, as in Sec. 6.1) as their acquisition function

$$\alpha_{\mathrm{PES}}(\xi;\mathcal{D}_t) = \mathcal{I}(\xi;\mathcal{D}_t) = H[p(\theta|\mathcal{D}_t)] - \mathbb{E}_{p(y|\xi,\mathcal{D}_t)}[H[p(\theta|\mathcal{D}_t \cup \{(\xi,y)\})]]. \tag{88}$$

However, the authors utilise the same insight as Houlsby et al. (2011) to write EIG in the equivalent form (see Proposition 9)

$$\mathcal{I}(\xi;\mathcal{D}_t) = H[p(y|\xi,\mathcal{D}_t)] - \mathbb{E}_{p(\theta|\mathcal{D}_t)}[H[p(y|\xi,\mathcal{D}_t,\theta)]]. \tag{89}$$

Within a GP model, the first term can be computed analytically, whilst the second is approximated by drawing samples of $\theta|\mathcal{D}_t$ and estimating $H[p(y|\xi,\mathcal{D}_t,\theta)]$ using expectation propagation (Minka, 2001). PES can be extended to batch acquisition in which we query f at multiple locations simultaneously on each iteration (Shah and Ghahramani, 2015).

In Maximum Entropy Search (MES) (Wang and Jegelka, 2017), the authors approach the problem differently. Instead of focusing on the latent variable of interest $\theta = \arg\max_{\xi\in\Xi} f(\xi)$, they instead formulate the problem with variable of interest $\theta_m = \max_{\xi\in\Xi} f(\xi)$. Here, θ_m is a one-dimensional random variable that represents the maximum *value* of the function f, rather than its arg max. The objective function for MES is then the EIG between a new observation y at design ξ and their parameter of interest θ_m

$$\alpha_{\mathrm{MES}}(\xi;\mathcal{D}_t) = H[p(y|\xi,\mathcal{D}_t)] - \mathbb{E}_{p(\theta_m|\mathcal{D}_t)}[H[p(y|\xi,\mathcal{D}_t,\theta_m)]]. \tag{90}$$

The MES objective may be easier to compute than PES with a GP model for f because θ_m is always one dimensional.

Computationally, a distinctive feature of BO with a GP model that sets it apart from computing the EIG in standard models (Sec. 4) is that many calculations can be performed analytically for the GP. For example, $H[p(y|\xi, \mathcal{D}_t)]$ is computed analytically in the PES acquisition function—this calculation would be intractable in a general model.

7.2.2 Other acquisition functions

Probability of improvement Perhaps the simplest acquisition rule, probability of improvement (Kushner, 1964) computes the probability that $f(\xi)$ is greater than some threshold τ

$$\alpha_{\mathrm{PI}}(\xi; \mathcal{D}_t) := \mathbb{P}(f(\xi) < \tau | \mathcal{D}_t). \tag{91}$$

Typically, the threshold τ is chosen adaptively to be the best objective value seen so far: $\tau_t = \max\{y_1, \ldots, y_t\}$.

Expected improvement A related acquisition rule is expected improvement (Mockus et al., 1978). This incorporates the amount by which the function value can be expected to increase at the location ξ, giving

$$\alpha_{\mathrm{EI}}(\xi; \mathcal{D}_t) := \mathbb{E}((f(\xi) - \tau)_+ | \mathcal{D}_t) \tag{92}$$

where $x_+ = \max(0, x)$.

Upper confidence bound Starting with theoretical work on multi-armed bandits (Lai and Robbins, 1985), upper confidence bound (UCB) acquisition rules have been popular. Srinivas et al. (2009) explicitly considered the application of UCB functions with a GP model for BO. In its most general form, we let $q_p(\cdot)$ denote to the p-quantile of a univariate distribution. Then the UCB-p acquisition function is

$$\alpha_{\text{UCB-}p}(\xi; \mathcal{D}_t) := q_p(f(\xi) | \mathcal{D}_t). \tag{93}$$

In the Gaussian case, $f(\xi)|\mathcal{D}_t \sim N(\mu(\xi|\mathcal{D}_t), \sigma(\xi|\mathcal{D}_t)^2)$, an equivalent parametrisation of the UCB acquisition function is

$$\alpha_{\text{UCB-}p}(\xi; \mathcal{D}_t) = \mu(\xi|\mathcal{D}_t) + \beta_p \sigma(\xi|\mathcal{D}_t) \tag{94}$$

where β_p is the p-quantile of the standard Normal distribution.

Thompson sampling Thompson (1933) proposed a *stochastic* acquisition rule for Bayesian optimisation. Given a sample of the functional posterior $f_t \sim p(f|\mathcal{D}_t)$, Thompson Sampling chooses the maximiser of this sample as the next sampling location. This amounts to using the acquisition function

$$\alpha_{\mathrm{TS}}(\xi; \mathcal{D}_t) = f_t(\xi) \text{ where } f_t \sim p(f|\mathcal{D}_t). \tag{95}$$

Hernández-Lobato et al. (2014) showed how the optimisation of a sample from the GP posterior can be approximately calculated.

8 Sequential Bayesian Experimental Design

We now lay out the theory of sequential experimentation more formally. Extending the basic sequential framework that we described for active learning in Sec. 5, we suppose that we have a sequence of T experiments. For each experiment, we pick ξ_t *adaptively* using the data that has already been observed[4]

[4]For an exchangeable model, the order of the data does not matter, so we could write $\mathcal{D}_{t-1} = \{(\xi_1, y_1), \ldots, (\xi_{t-1}, y_{t-1})\}$, which we implicitly assumed was the case in Sections 5 and 7. For a non-exchangeable model, we need to know the order that data was collected to conduct valid inference.

Algorithm 3 Terminal reward Sequential Bayesian Experimental Design

Require: Prior $p(\theta)$, model $p(y|\xi,\theta)$, initial data $\mathcal{D}_0$ may be empty.
  for step $t = 1, \ldots, T$ **do**
    Use policy to compute design $\xi_t \sim \pi(\xi|\mathcal{D}_{t-1})$
    Obtain experimental observation $y_t \sim p(y|\theta,\xi_t)$ with design ξ_t
    Set $\mathcal{D}_t = (\xi_1, y_1), \ldots, (\xi_t, y_t)$
  end for
  Choose decision δ to maximise $\mathbb{E}_{p(\theta|\mathcal{D}_T)}[U(\delta,\theta,\mathcal{D}_T)]$
  Obtain reward $U(\delta,\theta,\mathcal{D}_T)$

$\mathcal{D}_{t-1} = (\xi_1, y_1), \ldots, (\xi_{t-1}, y_{t-1})$. Given this design, we conduct an experiment using ξ_t and obtain outcome y_t. After each step of the experiment, our beliefs about θ are summarised by the posterior $p(\theta|\mathcal{D}_t)$, which is calculated as in Sec. 2.2. For an embedded model (Sec. 6), we would update our beliefs on the extended parameters ψ. For simplicity in this section, we assume we are not in an embedded model, unless otherwise stated, so the parameter θ is a full description of the model.

Policies and objectives The design ξ_t must be chosen on the basis of $\mathcal{D}_{t-1}$. A general abstraction to describe this is to introduce a *stochastic policy* $\pi(\xi|\mathcal{D}_{t-1})$ that maps from $\mathcal{D}_{t-1}$ to a distribution over designs. A special case of this is a deterministic policy, for which ξ_t is simply a function of $\mathcal{D}_{t-1}$.

In the sequential setting, it no longer makes sense to talk of the optimality of individual designs. Indeed, we cannot say whether a design ξ_2 will be optimal until we have observed the outcome y_1. Instead, we can describe optimality in terms of the *policy*—the policy which makes the best decision for ξ_2 for every possible value of y_1 would be an optimal policy.

Optimality also requires a criterion, so we must extend the utility-based approach of Sec. 3 based on Lindley (1972) to the sequential setting. Perhaps the most natural extension of Lindley's original theory, which is used implicitly by Huan and Marzouk (2016); Foster et al. (2021) is to consider a final decision δ which is taken after all data has been collected. In this *terminal reward* framework, we assume that we have a utility $U(\delta,\theta,\mathcal{D}_T)$. Once we have collected all our data, we make the decision to maximise $\mathbb{E}_{p(\theta|\mathcal{D}_T)}[U(\delta,\theta,\mathcal{D}_T)]$. The optimal policy, therefore, is the natural counterpart to equation (9), namely

$$\pi^* = \arg\max_{\pi} \mathbb{E}_{p(\mathcal{D}_T|\pi)}\left[\max_{\delta\in\Delta} \mathbb{E}_{p(\theta|\mathcal{D}_T)}[U(\delta,\theta,\mathcal{D}_T)]\right] \tag{96}$$

where $p(\mathcal{D}_T|\pi) = \mathbb{E}_{p(\theta)}\left[\prod_{t=1}^{T} \pi(\xi_t|\mathcal{D}_{t-1}) p(y_t|\theta,\xi_t)\right]$ for an exchangeable model[5]. The whole sequential experiment process is described in Algorithm 3.

Sequential EIG The natural extension of EIG to the sequential setting is to let (Foster et al., 2021)

$$U_{\mathcal{I}}(\mathcal{D}_T) = H[p(\theta)] - H[p(\theta|\mathcal{D}_T)] \tag{97}$$

or equivalently (Huan and Marzouk, 2016)

$$U_{\mathrm{KL}}(\mathcal{D}_T) = \mathrm{KL}[p(\theta|\mathcal{D}_T)\|p(\theta)]. \tag{98}$$

The intuition behind this utility is to reduce our uncertainty in the value of θ from the sum total of all our experiments. It is in the sequential setting that the naturalness of using information-theoretic objectives for experimental design becomes most apparent.

Example 12 (Shannon (1948); Lindley (1956))**.** *Consider a model with $\theta = (L, R)$ where L and R are discrete random variables with independent uniform priors $\theta \sim \mathrm{Unif}(n_L) \times \mathrm{Unif}(n_R)$. Suppose we have two*

[5]For a non-exchangeable model, it would be $p(\mathcal{D}_T|\pi) = \mathbb{E}_{p(\theta)}\left[\prod_{t=1}^{T} \pi(\xi_t|\mathcal{D}_{t-1}) p(y_t|\theta,\xi_t,\mathcal{D}_{t-1})\right]$

experimental designs at our disposal: ξ_L and ξ_R which produce noiseless outcomes giving the values of L and R respectively. Then, the utility of the sequence of experiments ξ_L, ξ_R is equal to sum of the utilities of the separate experiments ξ_L and ξ_R, i.e.

$$U_{\mathcal{I}}((\xi_L, L), (\xi_R, R)) = U_{\mathcal{I}}((\xi_L, L)) + U_{\mathcal{I}}((\xi_R, R)). \tag{99}$$

Proof. Direct calculation using equation (97) gives

$$U_{\mathcal{I}}((\xi_L, L), (\xi_R, R)) = \log(n_L n_R) = \log n_L + \log n_R \tag{100}$$
$$U_{\mathcal{I}}((\xi_L, L)) = \log n_L \tag{101}$$
$$U_{\mathcal{I}}((\xi_R, R)) = \log n_R. \tag{102}$$

□

Shannon (1948) showed that this property (along with other technical requirements) can only be satisfied by utilities based on entropy, making the EIG arguably the most natural criterion for sequential Bayesian experimental design.

Static and batch policies One simple approximation to the optimal policy is to select all designs $\xi_1, \ldots, \xi_T$ before the start of the experiment. This is known as *static* design, also called *open-loop* design (DiStefano III et al., 2014). In effect, static design collapses the sequential design problem back into the one-step problem of Sec. 3, albeit with a larger design space

$$\Xi^T = \{(\xi_1, \ldots, \xi_T) : \xi_t \in \Xi \text{ for all } t\} \tag{103}$$

and corresponding observation space. The probabilistic model is also augmented as in Sec. 2.2. Unfortunately, static design may be arbitrarily worse than the performance of the best fully adaptive policy.

Rather than choosing all T designs upfront, we could instead choose design in batches of B. There can be practical benefits for choosing designs in batches (Lyu et al., 2019), as opposed to choosing them individually. Mathematically, this *batch* design procedure fits back into the sequential theory we have already laid out, with batches of designs being chosen from the new design space Ξ^B. Static design corresponds to the case $B = T$ in which we stop after one batch. Batch design in active learning is discussed on page 19.

8.1 Greedy design policies

Designing a policy to solve equation (96) can be challenging. One common approximate strategy is *greedy* design (also called *myopic* design). A greedy policy can be characterised as choosing each design ξ_t assuming that this is the final experiment—i.e. that once ξ_t has been chosen and y_t observed, the sequence of experiments will terminate. This means that the greedy policy will choose ξ_t to maximise

$$\xi_t^* = \arg\max_{\xi \in \Xi} \mathbb{E}_{p(y|\xi_t, \mathcal{D}_{t-1})} \left[\max_{\delta \in \Delta} \mathbb{E}_{p(\theta|\mathcal{D}_t)}[U(\delta, \theta, \mathcal{D}_t)] \right]. \tag{104}$$

where $p(y|\xi_t, \mathcal{D}_{t-1}) = \mathbb{E}_{p(\theta|\mathcal{D}_{t-1})}[p(y|\theta, \xi)]$. We can see that this amounts to solving the one-step design optimisation problem of equation (9) at each t, with the important distinction that we *replace the original prior $p(\theta)$ with the posterior given existing data $p(\theta|\mathcal{D}_{t-1})$*. This agrees exactly with the greedy acquisition strategy described in Section 5.

There is a subtle distinction when θ is embedded in a larger model with parameters ψ (Sec. 6)—we must update our beliefs about all the parameters to $p(\psi|\mathcal{D}_{t-1})$ and use the predictive distribution $p(y|\xi_t, \mathcal{D}_{t-1}) = \mathbb{E}_{p(\psi|\mathcal{D}_{t-1})}[p(y|\psi, \theta)]$ for y. This agrees with the greedy acquisition strategy of Section 7, where we update the full model on the unknown function f at each step.

The greedy (myopic) approach to experimental design is very widely adopted (Cavagnaro et al., 2010; Drovandi et al., 2014; McGree et al., 2012; Myung et al., 2013; Foster et al., 2019). As noted, it is also

the typical sequential optimisation strategy in Bayesian active learning and Bayesian optimisation. One benefit of the greedy strategy is its simplicity—it effectively reduces the sequential experimental design problem to repeated applications of one-step design. It is typically observed that greedy optimisation for experimental design does not fail as catastrophically as greedy policies can do in general reinforcement learning tasks (Bakker et al., 2020). We explore a possible theoretical explanation for this phenomenon.

8.1.1 Submodularity

How much do we lose by using a greedy approach? A key theoretical tool for studying greedy policies is the notion of *submodularity* (Krause and Golovin, 2014). In short, if a utility function obeys submodularity (and a number of other conditions), then the theorem of Nemhauser et al. (1978) proves that a greedy strategy can achieve at least $(1 - 1/e) \approx 63\%$ of the best possible utility.

To precisely define submodularity, we must first define several other concepts. For any (finite) set V, the *power set* of V is $2^V = \{S : S \subseteq V\}$. A *set function* is any function $g : 2^V \to \mathbb{R}$. The *discrete derivative* of g is defined as

$$\Delta_g(e|S) = g(S \cup \{e\}) - g(S), \tag{105}$$

i.e. the extra value of adding element e to the set $S \subseteq V$. A set function $g : 2^V$ is *submodular* if, for every $A \subseteq B \subseteq V$ and for every $e \in V \setminus B$,

$$\Delta_g(e|B) \leq \Delta_g(e|A). \tag{106}$$

Intuitively, the value of adding e to the larger set B is smaller than the value of adding e to the smaller set A. Submodularity captures the intuitive notion of 'diminishing returns'. We also define a set function to be *monotone* if, for $A \subseteq B \subseteq V$, we have

$$g(A) \leq g(B). \tag{107}$$

The *greedy strategy* to maximise a monotone set function is to increment S by adding elements one at a time, following the rule

$$S_t = S_{t-1} \cup \{e_t\} \quad \text{where} \quad e_t = \arg\max_{e \in V} \Delta_g(e|S_{t-1}). \tag{108}$$

The following theorem of Nemhauser et al. (1978) shows that the greedy strategy performs near-optimally for submodular set functions.

Theorem 13 (Nemhauser et al. (1978))**.** *Let g be a monotone, submodular set function $g : 2^V \to \mathbb{R}$. Let $(S_t)_{t \geq 0}$ be obtained by the greedy strategy of equation (108). Then for any $t \leq |V|$ we have*

$$g(S_t) \geq (1 - 1/e) \max_{S \subseteq V, |S| = t} g(S). \tag{109}$$

This theorem proves that the greedy strategy can achieve at least $(1 - 1/e)$ of the best possible performance.

Submodularity for static experimental design There is a direct connection between the theory of submodularity and the *greedy construction of static experimental designs*. Indeed, the static experimental design problem is to choose $(\xi_1, \ldots, \xi_T)$ where each $\xi_t \in \Xi$. We assign a value to each static design following equation (96)

$$g(\xi_1, \ldots, \xi_T) = \mathbb{E}_{p(y_1, \ldots, y_T | \xi_1, \ldots, \xi_T)} \left[\max_{\delta \in \Delta} \mathbb{E}_{p(\theta|\mathcal{D}_T)}[U(\delta, \theta, \mathcal{D}_T)] \right]. \tag{110}$$

Suppose we could show that g is a monotone, submodular set function. Then the result of Theorem 13 would apply, meaning that we could construct a static design greedily by adding one element at a time.

To satisfy these conditions, we first need g to be invariant to the order of $\xi_1, \ldots, \xi_T$; we therefore assume that the model is exchangeable (Sec. 2.2). For the properties of submodularity and monotonicity, we need to choose a utility U, here we focus on the EIG with $U_{\mathcal{I}} = H[p(\theta)] - H[p(\theta|\mathcal{D}_T)]$. Then the function g becomes the mutual information between θ and $y_1, \ldots, y_T$ given $\xi_1, \ldots, \xi_T$.

Proposition 14 (Krause and Guestrin (2012)). *Suppose that, for any k and for designs $\xi_1, \ldots, \xi_k$, the random variables $y_1|\xi_1, \ldots, y_k|\xi_k$ are independent conditional on θ. Then the mutual information*

$$g(\{\xi_1, \ldots, \xi_k\}) = \mathbb{E}_{p(y_1,\ldots,y_k|\xi_1,\ldots,\xi_k)}\left[H[p(\theta)] - H[p(\theta|\mathcal{D}_k)]\right] \tag{111}$$

is a monontone, submodular set function.

The conditional independence assumption is equivalent to assuming an exchangeable model (Sec. 2.2) in which θ is the only model parameter (Sec. 6). Proposition 14 was also proved by Kirsch et al. (2019) in the context of BatchBALD for active learning.

8.1.2 Adaptive submodularity

The limitation of submodularity as a tool for analysing experimental design is that it does not consider *adaptive* design policies where the choice of a later design could be conditional on the outcome of earlier experiments. To address this limitation, Golovin and Krause (2011) introduced the notion of *adaptive submodularity*.

To define adaptive submodularity within our framework for experimental design, we focus on a discrete design space $|\Xi| < \infty$, and assume that there is no auxiliary decision δ. We can then define the *conditional expected marginal benefit* of a design ξ as

$$\Delta(\xi|\mathcal{D}_t) = \mathbb{E}_{p(\theta|\mathcal{D}_t)p(y|\theta,\xi)}[U(\theta, \mathcal{D}_t \cup \{(\xi, y)\}) - U(\theta, \mathcal{D}_t)]. \tag{112}$$

The utility U is *adaptive monotone* with respect to model $p(\theta)p(y|\theta, \xi)$ if the conditional expected marginal benefit of all designs is positive. That is, for all $t \geq 0, \mathcal{D}_t$ and $\xi \notin \mathcal{D}_t$ we have

$$\Delta(\xi|\mathcal{D}_t) \geq 0. \tag{113}$$

Furthermore, the utility U is *adaptive submodular* with respect to model $p(\theta)p(y|\theta, \xi)$ if for all $s \leq t$ and all nested datasets $\mathcal{D}_s \subseteq \mathcal{D}_t$ and for all designs $\xi \notin \mathcal{D}_t$ we have

$$\Delta(\xi|\mathcal{D}_t) \leq \Delta(\xi|\mathcal{D}_s). \tag{114}$$

This is a natural generalisation of submodularity for set functions, and again it captures the principle of 'diminishing returns'. Golovin and Krause (2011) were able to generalise the result of Nemhauser et al. (1978) to the adaptive case for noiseless experiments in which $p(y|\theta, \xi)$ is deterministic.

Theorem 15 (Golovin and Krause (2011)). *Let π^{greedy} be the greedy policy of equation (104). Assume U is adaptive monotone and adaptive submodular for model $p(\theta)p(y|\theta, \xi)$. Then,*

$$\mathbb{E}_{p(\mathcal{D}_T|\pi^{greedy})}\left[\mathbb{E}_{p(\theta|\mathcal{D}_T)}[U(\theta, \mathcal{D}_T)]\right] \geq (1 - 1/e) \sup_{\pi} \mathbb{E}_{p(\mathcal{D}_T|\pi)}\left[\mathbb{E}_{p(\theta|\mathcal{D}_T)}[U(\theta, \mathcal{D}_T)]\right]. \tag{115}$$

Golovin et al. (2010) explored the applicability of this framework to Bayesian active learning and Bayesian experimental design, focusing on the noiseless case in which $p(y|\theta, \xi)$ is deterministic. They proved that the information gain utility $U_{\mathcal{I}}$ is adaptive monotone and adaptive submodular, so the result of Theorem 15 applies in this case.

A key results of Chen et al. (2015) did away with the noiseless assumption. Instead, they assumed that different experimental outcomes are independent conditional on θ. This matches exactly with the factorisation equation (3). They also assume that θ takes finitely many values $|\Theta| < \infty$. The key bound is as follows

Theorem 16 (Theorem 2 of Chen et al. (2015)). *Let π^{greedy} be the adaptive greedy experimental design policy. Assume that observations y are conditionally independent given θ. Then, for any $\delta > 0$*

$$\mathbb{E}_{p(\mathcal{D}_T|\pi^{greedy})}\left[U_{\mathcal{I}}(\mathcal{D}_T)\right] \geq \left(1 - \exp\left[-\frac{1}{\gamma \max\{\log|\Theta|, \log(1/\delta)\}}\right]\right)\left(\sup_{\pi} \mathbb{E}_{p(\mathcal{D}_T|\pi)}\left[U_{\mathcal{I}}(\mathcal{D}_T)\right] - \delta\right) \tag{116}$$

where γ is a constant that depends on the noise distribution (see Chen et al. (2015)), and $U_{\mathcal{I}}$ is the information gain defined in equation (97).

Chen et al. (2017) went on to consider the case of noisy and correlated experimental outcomes (violating both the noiseless and the conditionally independent assumptions).

Finally, we note that the expected information gain is *not* adaptive submodular without assumption. This is elucidated by the following example, in which outcomes are not independent conditional on θ.

Example 17 (Inspired by Theorem 9 of Golovin et al. (2010))**.** *Consider a model with prior* $\theta \sim \text{Unif}(\{-1, 1\})$ *and with* $v \sim \text{Unif}(\{-1, 1\})$*. We have two potentially useful designs.* ξ_v *reports the value of* v*.* $\xi_{\theta v}$ *reports the value of* θv*. We also have* M *'dummy' designs* $\xi_1^d, \ldots, \xi_M^d$ *which report nothing. Clearly, the optimal strategy to learn* θ *is to conduct experiments with* ξ_v *and* $\xi_{\theta v}$ *in any order, since* $v \cdot \theta v = \theta v^2 = \theta$*. However, if we analyse a one-step optimal greedy strategy, we observe that every design apart from* $\xi_{\theta v}$ *is independent of the value of* θ*, and hence has EIG 0. We can also verify that, without knowing* v*, the posterior on* θ *given the outcome of design* $\xi_{\theta v}$ *is still* $\text{Unif}(\{-1, 1\})$*, hence the EIG of this design is also 0. Thus the greedy strategy will pick a design at random. If* M *is very large, the greedy strategy is likely to keep picking dummy designs.*

8.1.3 Asymptotic theory

A celebrated result of asymptotic statistics is the Bernstein–von Mises Theorem (Van der Vaart, 2000). In our experimental design set-up, this says that, under certain technical conditions and with i.i.d. random designs $\xi_t \overset{\text{i.i.d.}}{\sim} p(\xi)$, the posterior distribution $p(\theta|\mathcal{D}_t)$ is asymptotically Gaussian centred on the true value θ^* of the parameters of interest and with covariance matrix $t^{-1}M(\theta^*)^{-1}$. (Here, $M(\theta)$ is the Fisher information matrix, taking the expectation over $\xi \sim p(\xi)$.)

Paninski (2005) showed that a closely related result holds when designs are not random, but are chosen by greedy maximisation of the EIG.

Theorem 18 (Theorem 1 of Paninski (2005))**.** *Under certain technical conditions, the posterior distributions with greedy EIG maximisation are asymptotically Gaussian with mean* θ^* *and with covariance matrix* $t^{-1}\Sigma_{info}$*. Furthermore, if* $t^{-1}\Sigma_{iid}$ *is the asymptotic covariance with i.i.d. random designs, then*

$$\det \Sigma_{info} \leq \det \Sigma_{iid}. \tag{117}$$

This result tells us that the EIG maximisation strategy is no worse than i.i.d. sampling of designs, and that it will recover the true value θ^* in the limit as $t \to \infty$, i.e. the procedure is statistically consistent.

8.2 Non-greedy design policies

Whilst greedy policies enjoy computational tractability and some theoretical guarantees, a more direct approach to the problem of sequential experimental design is to seek the optimal policy that maximises equation (104). As we discuss in Sec. 9, finding this optimal policy can be cast in the language of reinforcement learning. In this section, we focus on computational approaches that have been suggested in the literature that specifically address non-greedy experimental design. These can generally be organised under two headings.

Forward sampling The forward sampling, or lookahead, family approaches relax the greedy assumption that the next experiment will be the last one. Instead, they assume that there will be m more experiments, and take account of these m future steps when deciding on the next experimental design. As m grows larger, this approach more closely approximates the truly optimal decision. However, with a larger m, the number of future outcomes to consider may grow exponentially. Such approaches either try to limit the number of outcomes considered, or else use a smaller value of m.

Backwards induction An alternative solution is to begin at the end. Classical optimisation theory (Bellman, 1966) shows that sequential optimisation problems are often more easily solved by starting with the final decision to be made at time T. For this final decision, the greedy solution is exactly optimal. The values of designs at later steps can be propagated backwards to inform earlier decisions. (See Sec. 9 for a fuller discussion.)

Non-greedy optimisation has typically been confined to low-dimensional cases within experimental design (Ryan et al., 2016). In medicine, Whitehead and Brunier (1995) and Whitehead and Williamson (1998) used a multi-step lookahead when finding optimal treatment doses. Berry and Ho (1988) explored optimal stopping when testing a one-sided hypothesis. Lewis and Berry (1994) applied backwards induction in a Bayesian clinical trials setting. Carlin et al. (1998) used forward sampling in a closely related clinical trial design problem. Brockwell and Kadane (2003) implemented backwards induction on a grid, and applied this to clinical trial planning. Müller et al. (2006) explored forward sampling for dose-response finding in clinical trials.

The main work to tackle the more general sequential experimental design problem, using the EIG utility, was Huan and Marzouk (2016). They used approximate dynamic programming to perform backwards induction by estimating the *value function*. The value function was then used to select optimal designs at each stage. The intermediate posterior distributions were estimated on a dynamically adapted grid.

In another line of work, González et al. (2016) build a predictor of future query locations given the current data. This allows them to use a forward sampling approach that is restricted to a single future trajectory. Jiang et al. (2020) use a related approach in which the future query locations are learned by repeatedly solving the *static* design optimisation problem with $T - t$ designs, but only using one of these designs at each step.

9 Bayesian Reinforcement Learning

Reinforcement learning (RL) (Sutton, 1990; Szepesvári, 2010) has a number of important and fascinating connections to sequential Bayesian experimental design. First, the problem of sequential experimental design is a reinforcement learning problem. Specifically, we will show how the set-up of the preceding section can be cast as a Bayes Adaptive Markov Decision Process (BAMDP) (Ross et al., 2007; Guez et al., 2012; Ghavamzadeh et al., 2016). Second, the problem of making sequential decision to learn about a model is deeply connected to *exploration* in model-based reinforcement learning (Sun et al., 2011; Shyam et al., 2019; Sekar et al., 2020).

9.1 Sequential Bayesian Experimental Design as a BAMDP

The BAMDP is a generalisation of the Markov Decision Process (Bellman, 1957; Duff, 2002) that accommodates an unknown transition model. Adopting the notation of Guez et al. (2012), a BAMDP can be described by its augmented state space S^+, action space A, augmented transition model $\mathcal{P}^+$, reward function R^+ and discount factor γ. The augmented state space consists of the *history* of all states and actions previously visited $h_t = s_1 a_1 \dots a_{t-1} s_t$. This data is used to update the transition model in a Bayesian manner, using

$$p(\mathcal{P}|h_t) \propto p(\mathcal{P}) p(h_t|\mathcal{P}). \tag{118}$$

For a sampled transition model, the probability of moving from s_t to s_{t+1} when action a_t was used is

$$p(s_{t+1}|s_t, a_t, \mathcal{P}) = \mathcal{P}(s_t, a_t, s_{t+1}). \tag{119}$$

The BAMDP transition model is therefore given by the marginal (Guez et al., 2012)

$$p(s_{t+1}|a_t, h_t) = \int p(\mathcal{P}|h_t) \mathcal{P}(s_t, a_t, s_{t+1}) \, d\mathcal{P}. \tag{120}$$

The reward for using action a in state s is sampled as $r \sim R(s, a)$. Planning in a BAMDP means finding the policy that maximises

$$\mathcal{J}(\pi) = \mathbb{E}_\pi \left[\sum_{t=1}^{T} \gamma^{-t} r_t \right]. \tag{121}$$

To set up sequential Bayesian experimental design in this framework, we associate the augmented history states with the data $\mathcal{D}_t$ up to time t. The actions of the BAMDP are the experimental designs ξ_t. The

transition model is associated with the model parameters θ (we assume in this section that we are not considering an embedded model). The 'transitions' of a sequential experiment are given by

$$p(\mathcal{D}_{t+1}|\mathcal{D}_t, \xi_{t+1}) = \mathbb{E}_{p(\theta|\mathcal{D}_t)}[p(y_{t+1}|\theta, \xi)] = \int p(\theta|\mathcal{D}_t) p(y_{t+1}|\theta, \xi_{t+1})\, d\theta \tag{122}$$

which agrees with equation (120) if we take $\mathcal{P}_\theta(y_{t+1}, \xi_{t+1}, y_t) = p(y_{t+1}|\theta, \xi_{t+1})$. Note that we write a_t as ξ_{t+1}, and that in the exchangeable experimental design case the transition model does not depend explicitly on y_t.

The only minor distinction from the set-up of Guez et al. (2012) is that the rewards in experimental design depend on the augmented state $\mathcal{D}_t$ rather than the state s_t. We can take the reward function for experimental design to be $R(\mathcal{D}_t) = \mathbf{1}[t = T] \max_{\delta \in \Delta} \mathbb{E}_{p(\theta|\mathcal{D}_t)}[U(\delta, \theta, \mathcal{D}_t)]$. Setting the discount factor $\gamma = 1$, we see that the BAMDP objective equation (121) is the same as the sequential experimental design problem equation (96). This shows the close connection between these two fields. For completeness, the value function and Q-function (Szepesvári, 2010) for Bayesian experimental design are given by

$$V^\pi(\mathcal{D}_t) = \mathbb{E}_{p(\mathcal{D}_T|\mathcal{D}_t, \pi)} \left[\max_{\delta \in \Delta} \mathbb{E}_{p(\theta|\mathcal{D}_T)}[U(\delta, \theta, \mathcal{D}_T)] \right] \tag{123}$$

$$Q^\pi(\mathcal{D}_t, \xi_{t+1}) = \mathbb{E}_{p(\mathcal{D}_T|\mathcal{D}_t, \xi_{t+1}, \pi)} \left[\max_{\delta \in \Delta} \mathbb{E}_{p(\theta|\mathcal{D}_T)}[U(\delta, \theta, \mathcal{D}_T)] \right] \tag{124}$$

where

$$p(\mathcal{D}_T|\mathcal{D}_t, \pi) = \mathbb{E}_{p(\theta|\mathcal{D}_t)} \left[\prod_{\tau=t+1}^{T} \pi(\xi_\tau|\mathcal{D}_{\tau-1}) p(y_\tau|\theta, \xi_\tau) \right] \tag{125}$$

$$p(\mathcal{D}_T|\mathcal{D}_t, \xi_{t+1}, \pi) = \mathbb{E}_{p(\theta|\mathcal{D}_t)} \left[p(y_{t+1}|\theta, \xi_{t+1}) \prod_{\tau=t+2}^{T} \pi(\xi_\tau|\mathcal{D}_{\tau-1}) p(y_\tau|\theta, \xi_\tau) \right]. \tag{126}$$

Belief states In the previous section, we followed Guez et al. (2012) and took the state space for experimental design to be the dataset $\mathcal{D}_t$. We see from equation (122) that the transition model only depends on $\mathcal{D}_t$ via the posterior $p(\theta|\mathcal{D}_t)$. Furthermore, our choice of reward function only depends on $p(\theta|\mathcal{D}_t)$ (plus an indicator that we have reached the final stage). Thus, it is sufficient to take $p(\theta|\mathcal{D}_t)$ as our augmented state. Posterior distributions treated as states are referred to as *belief states* (Igl et al., 2018). They have been utilised extensively in Bayesian RL (Igl et al., 2018; Zintgraf et al., 2019; Ghavamzadeh et al., 2016) and are beginning to be used in Bayesian experimental design (Huan and Marzouk, 2016).

9.2 Exploration

We have seen the close connection between sequential Bayesian experimental design and Bayesian RL. We associated the transition model of an unknown MDP with the model parameter θ. In this framing, we have a new interpretation of objective functions for experimental design—they encourage the collection of data to improve knowledge of the transition model and are motivated by model-derived quantities, rather than by an external reward signal. Utility functions for experimental design can thus be reinterpreted as rewards for *exploration* behaviour that leads to improved knowledge in a model of the environment.

The experimental design scenario is most closely connected with model-based reinforcement learning (Sutton, 1990). Specifically, we consider reinforcement learning settings in which we have a Bayesian parametric model of the environment with parameter θ. A range of authors have considered 'intrinsic rewards' (Singh et al., 2005)—unlike external rewards which are separate from the model and environment dynamics, intrinsic rewards encourage behaviour to learn about the environment. For example, Itti and Baldi (2006) used surprisal as an intrinsic reward—agents are encouraged to take actions for which the outcome is not predictable, and hence will be surprising. Mathematically, surprisal can be defined using predictive entropy. Empowerment (Klyubin et al., 2005; Salge et al., 2014; Mohamed and Rezende, 2015) is another intrinsic

reward signal that is based on conditional mutual information between state and action variables. Sajid et al. (2021) studied curiosity-driven exploration and the connection with free energy minimisation.

One line of research uses EIG as an intrinsic reward signal (Storck et al., 1995). This curiosity-driven exploration (Schmidhuber, 2010; Sun et al., 2011) is therefore the closest part of the RL literature to sequential experimental design. Specifically, Sun et al. (2011) utilise information gain as a reward. Given history h and h' such that h is a prefix of h' they define

$$\mathrm{IG}(h'\|h) = \mathrm{KL}\left(p(\theta|h')\|p(\theta|h)\right). \tag{127}$$

To motivate this choice, Sun et al. (2011) proved the following result (a more formal version of Example 12)

Proposition 19 (Sun et al. (2011))**.** *Let $h \subseteq h' \subseteq h''$ be histories such that h a prefix of h' and h' a prefix of h''. Suppose h' has been observed. Then,*

$$\mathbb{E}_{h''|h'}[\mathrm{IG}(h''\|h)] = \mathrm{IG}(h'\|h) + \mathbb{E}_{h''|h'}[\mathrm{IG}(h''\|h')] \tag{128}$$

so the information gain is additive in expectation.

Information gain for exploration was applied to robotics by Fung et al. (2016).

9.2.1 Computational approaches to exploration in Bayesian RL with EIG

To utilise information gain as an intrinsic reward for exploration requires approximation and optimisation of this quantity. Storck et al. (1995) focused on the tabular setting with finite states and actions, in which the transition model can be described with a finite number of parameters. Sun et al. (2011) also focused on the finite space case for their computations. Houthooft et al. (2016) tackled the continuous space problem. They used variational inference (Rezende et al., 2014; Kingma and Welling, 2014) to estimate the posterior distributions $p(\theta|\mathcal{D}_t)$. They then used the variational approximate posterior as a surrogate for the true posterior when computing the information gain reward. Information gain was combined with an external reward signal to balance exploration and exploitation. Shyam et al. (2019) used an ensemble to approximate the distribution $p(\theta|\mathcal{D}_t)$, to estimate information gain they replaced Shannon entropy with Rényi entropy which can be calculated for a mixture of Gaussians. Sekar et al. (2020) used a closely related approach. Rather than the Rényi entropy, they used the empirical variance of ensemble means as a way of estimating the intractable marginal entropy that occurs in the EIG.

References

Justin Alsing, Tom Charnock, Stephen Feeney, and Benjamin Wandelt. Fast likelihood-free cosmology with neural density estimators and active learning. *Monthly Notices of the Royal Astronomical Society*, 488(3): 4440–4458, 2019.

Billy Amzal, Frédéric Y Bois, Eric Parent, and Christian P Robert. Bayesian-optimal design via interacting particle systems. *Journal of the American Statistical association*, 101(474):773–785, 2006.

Christophe Andrieu, Nando De Freitas, Arnaud Doucet, and Michael I Jordan. An introduction to mcmc for machine learning. *Machine learning*, 50(1):5–43, 2003.

AC Atkinson and VV Fedorov. The design of experiments for discriminating between two rival models. *Biometrika*, 62(1):57–70, 1975.

Anthony Atkinson, Alexander Donev, and Randall Tobias. *Optimum experimental designs, with SAS*, volume 34. Oxford University Press, 2007.

Tim Bakker, Herke van Hoof, and Max Welling. Experimental design for mri by greedy policy search. *Advances in Neural Information Processing Systems*, 33, 2020.

Maximilian Balandat, Brian Karrer, Daniel Jiang, Samuel Daulton, Benjamin Letham, Andrew Gordon Wilson, and Eytan Bakshy. Botorch: A framework for efficient monte-carlo bayesian optimization. *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

George A Barnard, Gs M Jenkins, and CB Winsten. Likelihood inference and time series. *Journal of the Royal Statistical Society: Series A (General)*, 125(3):321–352, 1962.

Joakim Beck, Ben Mansour Dia, Luis FR Espath, Quan Long, and Raul Tempone. Fast bayesian experimental design: Laplace-based importance sampling for the expected information gain. *Computer Methods in Applied Mechanics and Engineering*, 334:523–553, 2018.

Richard Bellman. A markovian decision process. *Journal of mathematics and mechanics*, 6(5):679–684, 1957.

Richard Bellman. Dynamic programming. *Science*, 153(3731):34–37, 1966.

William H Beluch, Tim Genewein, Andreas Nürnberger, and Jan M Köhler. The power of ensembles for active learning in image classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 9368–9377, 2018.

James O Berger. *Statistical decision theory and Bayesian analysis*. Springer Science & Business Media, 2013.

James Bergstra, Rémi Bardenet, Yoshua Bengio, and Balázs Kégl. Algorithms for hyper-parameter optimization. *Advances in neural information processing systems*, 24, 2011.

José M Bernardo. Expected information as expected utility. *the Annals of Statistics*, pages 686–690, 1979.

Donald A Berry and Chih-Hsiang Ho. One-sided sequential stopping boundaries for clinical trials: A decision-theoretic approach. *Biometrics*, pages 219–227, 1988.

Concha Bielza, Peter Müller, and David Ríos Insua. Decision analysis by augmented probability simulation. *Management Science*, 45(7):995–1007, 1999.

Allan Birnbaum. On the foundations of statistical inference. *Journal of the American Statistical Association*, 57(298):269–306, 1962.

Benjamin Bloem-Reddy and Yee Whye Teh. Probabilistic symmetry and invariant neural networks. *arXiv preprint arXiv:1901.06082*, 2019.

George EP Box. Choice of response surface design and alphabetic optimality. Technical report, Wisconsin University-Madison Mathematics Research Center, 1982.

Johann Brehmer, Kyle Cranmer, Gilles Louppe, and Juan Pavez. Constraining effective field theories with machine learning. *Physical review letters*, 121(11):111801, 2018.

Anthony E Brockwell and Joseph B Kadane. A gridding method for bayesian sequential decision problems. *Journal of Computational and Graphical Statistics*, 12(3):566–584, 2003.

Sébastien Bubeck, Rémi Munos, and Gilles Stoltz. Pure exploration in multi-armed bandits problems. In *International conference on Algorithmic learning theory*, pages 23–37. Springer, 2009.

Alberto Giovanni Busetto, Cheng Soon Ong, and Joachim M Buhmann. Optimized expected information gain for nonlinear dynamical systems. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 97–104, 2009.

Roberto Calandra, Jan Peters, Carl Edward Rasmussen, and Marc Peter Deisenroth. Manifold gaussian processes for regression. In *2016 International Joint Conference on Neural Networks (IJCNN)*, pages 3338–3345. IEEE, 2016.

Bradley P Carlin, Joseph B Kadane, and Alan E Gelfand. Approaches for optimal sequential decision analysis in clinical trials. *Biometrics*, pages 964–975, 1998.

André Gustavo Carlon, Ben Mansour Dia, Luis Espath, Rafael Holdorf Lopez, and Raul Tempone. Nesterov-aided stochastic gradient methods using laplace approximation for bayesian design optimization. *Computer Methods in Applied Mechanics and Engineering*, 363:112909, 2020.

Daniel R Cavagnaro, Jay I Myung, Mark A Pitt, and Janne V Kujala. Adaptive design optimization: A mutual information-based approach to model discrimination in cognitive science. *Neural computation*, 22 (4):887–905, 2010.

Kathryn Chaloner and Isabella Verdinelli. Bayesian experimental design: A review. *Statistical Science*, pages 273–304, 1995.

Yuxin Chen, S Hamed Hassani, Amin Karbasi, and Andreas Krause. Sequential information maximization: When is greedy near-optimal? In *Conference on Learning Theory*, pages 338–363. PMLR, 2015.

Yuxin Chen, Hamed Hassani, and Andreas Krause. Near-optimal bayesian active learning with correlated and noisy tests. In *Artificial Intelligence and Statistics*, pages 223–231. PMLR, 2017.

Merlise A Clyde, Peter Müller, and Giovanni Parmigiani. Exploring expected utility surfaces by markov chains. Technical report, Institute of Statistics and Decision Sciences, Duke University, Durham, NC, USA, 1996.

Alex R Cook, Gavin J Gibson, and Christopher A Gilligan. Optimal observation times in experimental epidemic processes. *Biometrics*, 64(3):860–868, 2008.

Thomas M Cover. *Elements of information theory*. John Wiley & Sons, 1999.

David Roxbee Cox. *Principles of statistical inference*. Cambridge university press, 2006.

Katalin Csilléry, Michael GB Blum, Oscar E Gaggiotti, and Olivier François. Approximate bayesian computation (abc) in practice. *Trends in ecology & evolution*, 25(7):410–418, 2010.

Mahasen B Dehideniya, Christopher C Drovandi, and James M McGree. Optimal bayesian design for discriminating between models with intractable likelihoods in epidemiology. *Computational Statistics & Data Analysis*, 124:277–297, 2018.

Joseph J DiStefano III, Allen R Stubberud, and Ivan J Williams. *Schaum's outline of feedback and control systems*. McGraw-Hill Education, 2014.

Arnaud Doucet, Simon Godsill, and Christophe Andrieu. On sequential monte carlo sampling methods for bayesian filtering. *Statistics and computing*, 10(3):197–208, 2000.

Christopher C Drovandi and Anthony N Pettitt. Bayesian experimental design for models with intractable likelihoods. *Biometrics*, 69(4):937–948, 2013.

Christopher C Drovandi, James M McGree, and Anthony N Pettitt. A sequential monte carlo algorithm to incorporate model uncertainty in bayesian sequential design. *Journal of Computational and Graphical Statistics*, 23(1):3–24, 2014.

Michael O'Gordon Duff. *Optimal Learning: Computational procedures for Bayes-adaptive Markov decision processes*. University of Massachusetts Amherst, 2002.

Sergey Dushenko, Kapildeb Ambal, and Robert D McMichael. Sequential bayesian experiment design for optically detected magnetic resonance of nitrogen-vacancy centers. *Physical Review Applied*, 14(5):054036, 2020.

Agoston E Eiben, James E Smith, et al. *Introduction to evolutionary computing*, volume 53. Springer, 2003.

Gustav Elfving. Optimum allocation in linear regression theory. *The Annals of Mathematical Statistics*, pages 255–262, 1952.

Roger Fletcher. *Practical methods of optimization*. John Wiley & Sons, 2013.

Adam Foster, Martin Jankowiak, Elias Bingham, Paul Horsfall, Yee Whye Teh, Thomas Rainforth, and Noah Goodman. Variational Bayesian Optimal Experimental Design. In *Advances in Neural Information Processing Systems 32*, pages 14036–14047. Curran Associates, Inc., 2019.

Adam Foster, Martin Jankowiak, Matthew O'Meara, Yee Whye Teh, and Tom Rainforth. A unified stochastic gradient approach to designing bayesian-optimal experiments. In *International Conference on Artificial Intelligence and Statistics*, pages 2959–2969. PMLR, 2020.

Adam Foster, Desi R Ivanova, Ilyas Malik, and Tom Rainforth. Deep adaptive design: Amortizing sequential bayesian experimental design. *arXiv preprint arXiv:2103.02438*, 2021.

Linton C Freeman. *Elementary applied statistics: for students in behavioral science*. New York: Wiley, 1965.

Nicholas C Fung, Carlos Nieto-Granda, Jason M Gregory, and John G Rogers. Autonomous exploration using an information gain metric. Technical report, US Army Research Laboratory Adelphi United States, 2016.

Yarin Gal and Zoubin Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, pages 1050–1059. PMLR, 2016.

Yarin Gal, Riashat Islam, and Zoubin Ghahramani. Deep bayesian active learning with image data. In *International Conference on Machine Learning*, pages 1183–1192. PMLR, 2017.

Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. *Bayesian data analysis*. Chapman and Hall/CRC, 2013.

Mohammad Ghavamzadeh, Shie Mannor, Joelle Pineau, and Aviv Tamar. Bayesian reinforcement learning: A survey. *arXiv preprint arXiv:1609.04436*, 2016.

Daniel Golovin and Andreas Krause. Adaptive submodularity: Theory and applications in active learning and stochastic optimization. *Journal of Artificial Intelligence Research*, 42:427–486, 2011.

Daniel Golovin, Andreas Krause, and Debajyoti Ray. Near-optimal bayesian active learning with noisy observations. In *Advances in Neural Information Processing Systems*, pages 766–774, 2010.

Pedro J Gonçalves, Jan-Matthis Lueckmann, Michael Deistler, Marcel Nonnenmacher, Kaan Öcal, Giacomo Bassetto, Chaitanya Chintaluri, William F Podlaski, Sara A Haddad, Tim P Vogels, et al. Training deep neural density estimators to identify mechanistic models of neural dynamics. *Elife*, 9:e56261, 2020.

Wenbo Gong, Sebastian Tschiatschek, Richard Turner, Sebastian Nowozin, José Miguel Hernández-Lobato, and Cheng Zhang. Icebreaker: Element-wise active information acquisition with bayesian deep latent gaussian model. *arXiv preprint arXiv:1908.04537*, 2019.

Javier González, Michael Osborne, and Neil Lawrence. Glasses: Relieving the myopia of bayesian optimisation. In *Artificial Intelligence and Statistics*, pages 790–799. PMLR, 2016.

Arthur Guez, David Silver, and Peter Dayan. Efficient bayes-adaptive reinforcement learning using sample-based search. In *Advances in neural information processing systems*, pages 1025–1033, 2012.

Markus Hainy, Werner G Müller, and Helga Wagner. Likelihood-free simulation-based optimal design with an application to spatial extremes. *Stochastic Environmental Research and Risk Assessment*, 30(2):481–492, 2016.

M Hamada, HF Martz, CS Reese, and AG Wilson. Finding near-optimal bayesian experimental designs via genetic algorithms. *The American Statistician*, 55(3):175–181, 2001.

Cong Han and Kathryn Chaloner. Bayesian experimental design for nonlinear mixed-effects models with application to hiv dynamics. *Biometrics*, 60(1):25–33, 2004.

Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning (2nd edition)*. Springer-Verlag, 2009.

W Keith Hastings. *Monte Carlo sampling methods using Markov chains and their applications*. Oxford University Press, 1970.

Daniel W Heck and Edgar Erdfelder. Maximizing the expected information gain of cognitive modeling via design optimization. *Computational Brain & Behavior*, 2(3):202–209, 2019.

Frank Heinrich, Paul A Kienzle, David P Hoogerheide, and Mathias Lösche. Information gain from isotopic contrast variation in neutron reflectometry on protein–membrane complex structures. *Journal of applied crystallography*, 53(3):800–810, 2020.

Philipp Hennig and Christian J Schuler. Entropy search for information-efficient global optimization. *Journal of Machine Learning Research*, 13(Jun):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, 2014.

Matthew Hoffman, Bobak Shahriari, and Nando Freitas. On correlation and budget constraints in model-based bandit optimization with application to automatic machine learning. In *Artificial Intelligence and Statistics*, pages 365–374. PMLR, 2014.

Neil Houlsby, Ferenc Huszár, Zoubin Ghahramani, and Máté Lengyel. Bayesian active learning for classification and preference learning. *arXiv preprint arXiv:1112.5745*, 2011.

Rein Houthooft, Xi Chen, Yan Duan, John Schulman, Filip De Turck, and Pieter Abbeel. Vime: Variational information maximizing exploration. *arXiv preprint arXiv:1605.09674*, 2016.

Xun Huan and Youssef Marzouk. Gradient-based stochastic optimization methods in bayesian experimental design. *International Journal for Uncertainty Quantification*, 4(6), 2014.

Xun Huan and Youssef M Marzouk. Simulation-based optimal bayesian experimental design for nonlinear systems. *Journal of Computational Physics*, 232(1):288–317, 2013.

Xun Huan and Youssef M Marzouk. Sequential bayesian optimal experimental design via approximate dynamic programming. *arXiv preprint arXiv:1604.08320*, 2016.

Frank Hutter, Holger Hoos, and Kevin Leyton-Brown. An evaluation of sequential model-based optimization for expensive blackbox functions. In *Proceedings of the 15th annual conference companion on Genetic and evolutionary computation*, pages 1209–1216, 2013.

Maximilian Igl, Luisa Zintgraf, Tuan Anh Le, Frank Wood, and Shimon Whiteson. Deep variational reinforcement learning for pomdps. In *International Conference on Machine Learning*, pages 2117–2126. PMLR, 2018.

Laurent Itti and Pierre F Baldi. Bayesian surprise attracts human attention. In *Advances in neural information processing systems*, pages 547–554. Citeseer, 2006.

Shali Jiang, Henry Chai, Javier Gonzalez, and Roman Garnett. Binoculars for efficient, nonmyopic sequential experimental design. In *International Conference on Machine Learning*, pages 4794–4803. PMLR, 2020.

Ashish Kapoor, Kristen Grauman, Raquel Urtasun, and Trevor Darrell. Active learning with gaussian processes for object categorization. In *2007 IEEE 11th International Conference on Computer Vision*, pages 1–8. IEEE, 2007.

Alex Kendall, Vijay Badrinarayanan, and Roberto Cipolla. Bayesian segnet: Model uncertainty in deep convolutional encoder-decoder architectures for scene understanding. *arXiv preprint arXiv:1511.02680*, 2015.

Jack Kiefer and Jacob Wolfowitz. Optimum designs in regression problems. *The annals of mathematical statistics*, pages 271–294, 1959.

Diederik P Kingma and Max Welling. Auto-encoding variational Bayes. In *ICLR*, 2014.

Andreas Kirsch, Joost Van Amersfoort, and Yarin Gal. Batchbald: Efficient and diverse batch acquisition for deep bayesian active learning. *Advances in neural information processing systems*, 32:7026–7037, 2019.

Steven Kleinegesse and Michael Gutmann. Efficient Bayesian experimental design for implicit models. *arXiv preprint arXiv:1810.09912*, 2018.

Alexander S Klyubin, Daniel Polani, and Chrystopher L Nehaniv. All else being equal be empowered. In *European Conference on Artificial Life*, pages 744–753. Springer, 2005.

Andreas Krause and Daniel Golovin. Submodular function maximization. *Tractability*, 3:71–104, 2014.

Andreas Krause and Carlos E Guestrin. Near-optimal nonmyopic value of information in graphical models. *arXiv preprint arXiv:1207.1394*, 2012.

John Kruschke. *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Academic Press, 2014.

Hendrik Kuck, Nando de Freitas, and Arnaud Doucet. Smc samplers for bayesian optimal nonlinear design. In *2006 IEEE Nonlinear Statistical Signal Processing Workshop*, pages 99–102. IEEE, 2006.

Harold J Kushner. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *J of Basic Engineering*, 1964.

Tze Leung Lai and Herbert Robbins. Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, 6(1):4–22, 1985.

Miguel Lázaro-Gredilla, Joaquin Quinonero-Candela, Carl Edward Rasmussen, and Aníbal R Figueiras-Vidal. Sparse spectrum gaussian process regression. *The Journal of Machine Learning Research*, 11: 1865–1881, 2010.

Jeremy Lewi, Robert Butera, and Liam Paninski. Sequential optimal design of neurophysiology experiments. *Neural Computation*, 21(3):619–687, 2009.

David D Lewis and William A Gale. A sequential algorithm for training text classifiers. In *SIGIR'94*, pages 3–12. Springer, 1994.

Roger J Lewis and Donald A Berry. Group sequential clinical trials: a classical evaluation of bayesian decision-theoretic designs. *Journal of the American Statistical Association*, 89(428):1528–1534, 1994.

Dennis V Lindley. On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, pages 986–1005, 1956.

Dennis V Lindley. *Bayesian statistics, a review*, volume 2. SIAM, 1972.

Quan Long. Multimodal information gain in bayesian design of experiments. *Computational Statistics*, pages 1–21, 2021.

Quan Long, Marco Scavino, Raúl Tempone, and Suojin Wang. Fast estimation of expected information gains for Bayesian experimental designs based on Laplace approximations. *Computer Methods in Applied Mechanics and Engineering*, 259:24–39, 2013.

Thomas J Loredo. Bayesian adaptive exploration. In *AIP Conference Proceedings*, volume 707, pages 330–346. American Institute of Physics, 2004.

Jiankun Lyu, Sheng Wang, Trent E Balius, Isha Singh, Anat Levit, Yurii S Moroz, Matthew J O'Meara, Tao Che, Enkhjargal Algaa, Kateryna Tolmachova, et al. Ultra-large library docking for discovering new chemotypes. *Nature*, 566(7743):224, 2019.

Chao Ma, Sebastian Tschiatschek, Konstantina Palla, José Miguel Hernández-Lobato, Sebastian Nowozin, and Cheng Zhang. EDDI: Efficient dynamic discovery of high-value information with Partial VAE. *arXiv preprint arXiv:1809.11142*, 2018.

David JC MacKay. Information-based objective functions for active data selection. *Neural computation*, 4 (4):590–604, 1992.

James Matthew McGree, Christopher C Drovandi, MH Thompson, JA Eccleston, SB Duffull, Kerrie Mengersen, Anthony N Pettitt, and Timothy Goggin. Adaptive bayesian compound designs for dose finding studies. *Journal of Statistical Planning and Inference*, 142(6):1480–1492, 2012.

Ruth K Meyer and Christopher J Nachtsheim. The coordinate-exchange algorithm for constructing exact optimal experimental designs. *Technometrics*, 37(1):60–69, 1995.

Thomas Peter Minka. *A family of algorithms for approximate Bayesian inference*. PhD thesis, Massachusetts Institute of Technology, 2001.

Jonas Mockus, Vytautas Tiesis, and Antanas Zilinskas. The application of bayesian methods for seeking the extremum. *Towards global optimization*, 2(117-129):2, 1978.

Shakir Mohamed and Danilo Jimenez Rezende. Variational information maximisation for intrinsically motivated reinforcement learning. *arXiv preprint arXiv:1509.08731*, 2015.

Peter Müller. Simulation based optimal design. *Handbook of Statistics*, 25:509–518, 2005.

Peter Müller and Giovanni Parmigiani. Optimal design via curve fitting of monte carlo experiments. *Journal of the American Statistical Association*, 90(432):1322–1330, 1995.

Peter Müller, Bruno Sansó, and Maria De Iorio. Optimal bayesian design by inhomogeneous markov chain simulation. *Journal of the American Statistical Association*, 99(467):788–798, 2004.

Peter Müller, Don A Berry, Andrew P Grieve, and Michael Krams. A bayesian decision-theoretic dose-finding trial. *Decision analysis*, 3(4):197–207, 2006.

Jay I Myung, Daniel R Cavagnaro, and Mark A Pitt. A tutorial on adaptive design optimization. *Journal of mathematical psychology*, 57(3-4):53–67, 2013.

Willie Neiswanger, Kirthevasan Kandasamy, Barnabas Poczos, Jeff Schneider, and Eric Xing. Probo: Versatile bayesian optimization using any probabilistic programming language. *arXiv preprint arXiv:1901.11515*, 2019.

John A Nelder and Roger Mead. A simplex method for function minimization. *The computer journal*, 7(4): 308–313, 1965.

George L Nemhauser, Laurence A Wolsey, and Marshall L Fisher. An analysis of approximations for maximizing submodular set functions—i. *Mathematical programming*, 14(1):265–294, 1978.

James Neufeld, Andras Gyorgy, Csaba Szepesvári, and Dale Schuurmans. Adaptive monte carlo via bandit allocation. In *International Conference on Machine Learning*, pages 1944–1952. PMLR, 2014.

Bernt Øksendal. Stochastic differential equations. In *Stochastic differential equations*, pages 65–84. Springer, 2003.

Michael A Osborne, Roman Garnett, and Stephen J Roberts. Gaussian processes for global optimization. In *3rd international conference on learning and intelligent optimization (LION3)*, pages 1–15. Citeseer, 2009.

Antony Overstall and James McGree. Bayesian design of experiments for intractable likelihood models using coupled auxiliary models and multivariate emulation. *Bayesian Analysis*, 15(1):103–131, 2020.

Antony M Overstall and David C Woods. Bayesian design of experiments using approximate coordinate exchange. *Technometrics*, 59(4):458–470, 2017.

J Lynn Palmer and Peter Müller. Bayesian optimal design in population models for haematologic data. *Statistics in medicine*, 17(14):1613–1622, 1998.

Liam Paninski. Asymptotic theory of information-theoretic experimental design. *Neural Computation*, 17 (7):1480–1507, 2005.

Costas Papadimitriou. Optimal sensor placement methodology for parametric identification of structural systems. *Journal of sound and vibration*, 278(4-5):923–947, 2004.

Andy Pole, Mike West, and Jeff Harrison. *Applied Bayesian forecasting and time series analysis*. Chapman and Hall/CRC, 2018.

Remus Pop and Patric Fulop. Deep ensemble bayesian active learning: Addressing the mode collapse issue in monte carlo dropout via ensembles. *arXiv preprint arXiv:1811.03897*, 2018.

David J Price, Nigel G Bean, Joshua V Ross, and Jonathan Tuke. On the efficient determination of optimal bayesian experimental designs using abc: A case study in optimal observation of epidemics. *Journal of Statistical Planning and Inference*, 172:1–15, 2016.

David J Price, Nigel G Bean, Joshua V Ross, and Jonathan Tuke. An induced natural selection heuristic for finding optimal bayesian experimental designs. *Computational Statistics & Data Analysis*, 126:112–124, 2018.

Joaquin Quinonero-Candela and Carl Edward Rasmussen. A unifying view of sparse approximate gaussian process regression. *The Journal of Machine Learning Research*, 6:1939–1959, 2005.

Tom Rainforth. *Automating Inference, Learning, and Design using Probabilistic Programming*. PhD thesis, University of Oxford, 2017.

Tom Rainforth, Rob Cornish, Hongseok Yang, Andrew Warrington, and Frank Wood. On nesting monte carlo estimators. In *International Conference on Machine Learning*, pages 4267–4276. PMLR, 2018.

Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra. Stochastic backpropagation and approximate inference in deep generative models. In *Proceedings of the 31st International Conference on Machine Learning*, volume 32, pages 1278–1286, 2014.

Herbert Robbins and Sutton Monro. A stochastic approximation method. *The annals of mathematical statistics*, pages 400–407, 1951.

Christian Robert. *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Springer Science & Business Media, 2007.

Stephane Ross, Brahim Chaib-draa, and Joelle Pineau. Bayes-adaptive pomdps. In *NIPS*, pages 1225–1232, 2007.

N Roy and A McCallum. Toward optimal active learning through sampling estimation of error reduction. int. conf. on machine learning, 2001.

Daniel Russo and Benjamin Van Roy. Learning to optimize via posterior sampling. *Mathematics of Operations Research*, 39(4):1221–1243, 2014.

Elizabeth G. Ryan, Christopher C. Drovandi, M. Helen Thompson, and Anthony N. Pettitt. Towards bayesian experimental design for nonlinear models that require a large number of sampling times. *Computational Statistics & Data Analysis*, 70:45–60, 2014. ISSN 0167-9473. doi: https://doi.org/10.1016/j.csda.2013.08.017. URL `https://www.sciencedirect.com/science/article/pii/S0167947313003149`.

Elizabeth G Ryan, Christopher C Drovandi, and Anthony N Pettitt. Fully Bayesian experimental design for pharmacokinetic studies. *Entropy*, 17(3):1063–1089, 2015.

Elizabeth G Ryan, Christopher C Drovandi, James M McGree, and Anthony N Pettitt. A review of modern computational algorithms for bayesian optimal design. *International Statistical Review*, 84(1):128–154, 2016.

Kenneth J Ryan. Estimating expected information gains for experimental designs with application to the random fatigue-limit model. *Journal of Computational and Graphical Statistics*, 12(3):585–603, 2003.

Noor Sajid, Philip J Ball, Thomas Parr, and Karl J Friston. Active inference: demystified and compared. *Neural Computation*, 33(3):674–712, 2021.

Christoph Salge, Cornelius Glackin, and Daniel Polani. Empowerment–an introduction. In *Guided Self-Organization: Inception*, pages 67–114. Springer, 2014.

Tobias Scheffer, Christian Decomain, and Stefan Wrobel. Active hidden markov models for information extraction. In *International Symposium on Intelligent Data Analysis*, pages 309–318. Springer, 2001.

Jürgen Schmidhuber. Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development*, 2(3):230–247, 2010.

Paola Sebastiani and Henry P Wynn. Maximum entropy sampling and optimal Bayesian experimental design. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(1), 2000.

Ramanan Sekar, Oleh Rybkin, Kostas Daniilidis, Pieter Abbeel, Danijar Hafner, and Deepak Pathak. Planning to explore via self-supervised world models. In *International Conference on Machine Learning*, pages 8583–8592. PMLR, 2020.

Burr Settles. *Active learning literature survey*. PhD thesis, University of Wisconsin-Madison Department of Computer Sciences, 2009.

Burr Settles and Mark Craven. An analysis of active learning strategies for sequence labeling tasks. In *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, pages 1070–1079, 2008.

Ben Shababo, Brooks Paige, Ari Pakman, and Liam Paninski. Bayesian inference and online experimental design for mapping neural microcircuits. In *Advances in Neural Information Processing Systems*, pages 1304–1312, 2013.

Amar Shah and Zoubin Ghahramani. Parallel predictive entropy search for batch global optimization of expensive objective functions. *arXiv preprint arXiv:1511.07130*, 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, 2015.

Claude Elwood Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27 (3):379–423, 1948.

Pranav Shyam, Wojciech Jaśkowski, and Faustino Gomez. Model-based active exploration. In *International conference on machine learning*, pages 5779–5788. PMLR, 2019.

Satinder Singh, Andrew G Barto, and Nuttapong Chentanez. Intrinsically motivated reinforcement learning. Technical report, MASSACHUSETTS UNIV AMHERST DEPT OF COMPUTER SCIENCE, 2005.

Scott A Sisson, Yanan Fan, and Mark Beaumont. *Handbook of approximate Bayesian computation*. CRC Press, 2018.

Lewis Smith and Yarin Gal. Understanding measures of uncertainty for adversarial example detection. *arXiv preprint arXiv:1803.08533*, 2018.

Edward Snelson and Zoubin Ghahramani. Sparse gaussian processes using pseudo-inputs. *Advances in neural information processing systems*, 18:1257, 2006.

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.

Jasper Snoek, Oren Rippel, Kevin Swersky, Ryan Kiros, Nadathur Satish, Narayanan Sundaram, Mostofa Patwary, Mr Prabhat, and Ryan Adams. Scalable bayesian optimization using deep neural networks. In *International conference on machine learning*, pages 2171–2180. PMLR, 2015.

James C Spall. An overview of the simultaneous perturbation method for efficient optimization. *Johns Hopkins apl technical digest*, 19(4):482–492, 1998.

Niranjan Srinivas, Andreas Krause, Sham M Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv preprint arXiv:0912.3995*, 2009.

Jan Storck, Sepp Hochreiter, Jürgen Schmidhuber, et al. Reinforcement driven information acquisition in non-deterministic environments. In *Proceedings of the international conference on artificial neural networks, Paris*, volume 2, pages 159–164. Citeseer, 1995.

Jonathan R Stroud, Peter Müller, and Gary L Rosner. Optimal sampling times in population pharmacokinetic studies. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 50(3):345–359, 2001.

Yi Sun, Faustino Gomez, and Jürgen Schmidhuber. Planning to be surprised: Optimal bayesian exploration in dynamic environments. In *International Conference on Artificial General Intelligence*, pages 41–51. Springer, 2011.

Richard S Sutton. Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. In *Machine learning proceedings 1990*, pages 216–224. Elsevier, 1990.

Csaba Szepesvári. Algorithms for reinforcement learning. *Synthesis lectures on artificial intelligence and machine learning*, 4(1):1–103, 2010.

Gabriel Terejanu, Rochan R Upadhyay, and Kenji Miki. Bayesian experimental design for the active nitridation of graphite by atomic nitrogen. *Experimental Thermal and Fluid Science*, 36:178–193, 2012.

Owen Thomas, Ritabrata Dutta, Jukka Corander, Samuel Kaski, and Michael U Gutmann. Likelihood-free inference by ratio estimation. *arXiv preprint arXiv:1611.10242*, 2016.

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.

Robert K Tsutakawa. Design of experiment for bioassay. *Journal of the American Statistical Association*, 67(339):584–590, 1972.

Jojanneke van den Berg, Andrew Curtis, and Jeannot Trampert. Optimal nonlinear bayesian experimental design: an application to amplitude versus offset experiments. *Geophysical Journal International*, 155(2): 411–421, 2003.

Aad W Van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.

Peter JM Van Laarhoven and Emile HL Aarts. Simulated annealing. In *Simulated annealing: Theory and applications*, pages 7–15. Springer, 1987.

Joep Vanlier, Christian A Tiemann, Peter AJ Hilbers, and Natal AW van Riel. A Bayesian approach to targeted experiment design. *Bioinformatics*, 28(8):1136–1142, 2012.

Joep Vanlier, Christian A Tiemann, Peter AJ Hilbers, and Natal AW van Riel. Optimal experiment design for model selection in biochemical networks. *BMC systems biology*, 8(1):1–16, 2014.

Julien Villemonteix, Emmanuel Vazquez, and Eric Walter. An informational approach to the global optimization of expensive-to-evaluate functions. *Journal of Global Optimization*, 44(4):509–534, 2009.

Benjamin T Vincent and Tom Rainforth. The DARC toolbox: automated, flexible, and efficient delayed and risky choice experiments using bayesian adaptive design. *Retrieved from psyarxiv.com/yehjb*, 2017.

Julius von Kügelgen, Paul K Rubenstein, Bernhard Schölkopf, and Adrian Weller. Optimal experimental design via bayesian optimization: active causal structure learning for gaussian process networks. *arXiv preprint arXiv:1910.03962*, 2019.

Jon Wakefield. An expected loss approach to the design of dosage regimens via sampling-based methods. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 43(1):13–29, 1994.

Zi Wang and Stefanie Jegelka. Max-value entropy search for efficient bayesian optimization. In *International Conference on Machine Learning*, pages 3627–3635. PMLR, 2017.

Andrew B Watson. Quest+: A general multidimensional bayesian adaptive psychometric method. *Journal of Vision*, 17(3):10–10, 2017.

John Whitehead and Hazel Brunier. Bayesian decision procedures for dose determining experiments. *Statistics in medicine*, 14(9):885–893, 1995.

John Whitehead and David Williamson. Bayesian decision procedures based on logistic regression models for dose-finding studies. *Journal of Biopharmaceutical Statistics*, 8(3):445–467, 1998.

Christopher K Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA, 2006.

Liu Yang, Rong Jin, and Rahul Sukthankar. Bayesian active distance metric learning. *arXiv preprint arXiv:1206.5283*, 2012.

Sue Zheng, Jason Pacheco, and John Fisher. A robust approach to sequential information theoretic planning. In *International Conference on Machine Learning*, pages 5941–5949, 2018.

Luisa Zintgraf, Kyriacos Shiarlis, Maximilian Igl, Sebastian Schulze, Yarin Gal, Katja Hofmann, and Shimon Whiteson. Varibad: A very good method for bayes-adaptive deep rl via meta-learning. *arXiv preprint arXiv:1910.08348*, 2019.