



Online Simulator-Based Experimental Design for Cognitive Model Selection

Alexander Aushev¹ · Aini Putkonen² · Grégoire Clarté³ · Suyog Chandramouli² · Luigi Acerbi³ · Samuel Kaski^{1,4} · Andrew Howes⁵

Accepted: 18 July 2023
© The Author(s) 2023

Abstract

The problem of model selection with a limited number of experimental trials has received considerable attention in cognitive science, where the role of experiments is to discriminate between theories expressed as computational models. Research on this subject has mostly been restricted to optimal experiment design with analytically tractable models. However, cognitive models of increasing complexity with intractable likelihoods are becoming more commonplace. In this paper, we propose BOSMOS, an approach to experimental design that can select between computational models without tractable likelihoods. It does so in a data-efficient manner by sequentially and adaptively generating informative experiments. In contrast to previous approaches, we introduce a novel simulator-based utility objective for design selection and a new approximation of the model likelihood for model selection. In simulated experiments, we demonstrate that the proposed BOSMOS technique can accurately select models in up to two orders of magnitude less time than existing LFI alternatives for three cognitive science tasks: memory retention, sequential signal detection, and risky choice.

Keywords Model selection · Experimental design · Likelihood-free inference · Cognitive models

Introduction

The problem of selecting between competing models of cognition is critical to progress in cognitive science. The goal of model selection is to choose the model that most closely represents the cognitive process that generated the observed behavioural data. Typically, model selection involves maximising the fit of each model's parameters to the data and balancing the quality of the model fit with its complexity. It is crucial that any model selection method used is robust and sample-efficient and that it correctly measures how well each model approximates the data-generating cognitive process.

It is also crucial that any model selection process is provided with high-quality data from well-designed experiments and that these data are sufficiently informative to support efficient selection. Research on *optimal experimental design* (OED) addresses this problem by focusing on how to design

experiments that support parameter estimation of single models and, in some cases, maximise information for model selection (Cavagnaro et al., 2010; Moon et al., 2022; Blau et al., 2022).

However, one outstanding difficulty in model selection is that many models do not have tractable likelihoods. The model likelihoods represent the probability of observed data being produced by model parameters and simplify tractable inference. In their absence, *likelihood-free inference* (LFI) methods can be used, which rely on forward simulations (or samples from the model) to replace the likelihood. Another difficulty is that existing methods for OED are slow—very slow—which makes them impractical for many applications. In this paper, we address these problems by investigating a new algorithm that automatically and adaptively designs experiments for likelihood-free models much more quickly than previous approaches. The new algorithm is called *Bayesian optimisation for simulator-based model selection* (BOSMOS).

In BOSMOS, model selection is conducted in a Bayesian framework. In this setting, inference is carried out using *marginal likelihood*, which incorporates, by definition, a penalty for model complexity, i.e., Occam's Razor. Addi-

✉ Alexander Aushev
alexander.aushev@aalto.fi

✉ Andrew Howes
a.howes@bham.ac.uk

Extended author information available on the last page of the article

tionally, the Bayesian framework allows getting Bayesian posteriors over all possible values rather than point estimates; this is crucial for quantifying uncertainty, for instance, when multiple models can explain the data similarly well (non-identifiability or poor identifiability; (Anderson, 1978; Acerbi et al., 2014)) or when some of the models are *misspecified* (Lee et al., 2019), i.e., the model makes overly simplified or incorrect assumptions about the behaviour. However, it is important to acknowledge that no OED method, including BOSMOS, can guarantee a correct solution in instances of significant model misfit, as there is no clear-cut theoretical solution to this complex issue, which would likely necessitate a more nuanced modelling of human behaviour. These problems are further exacerbated in computational cognitive modelling, where non-identifiability also arises due to human strategic flexibility (Howes et al., 2009; Madsen et al., 2019; Kangasrääsiö et al., 2019; Oulasvirta et al., 2022). For these reasons, there is an interest in Bayesian approaches in computational cognitive science (Overstall & Woods, 2017; Tauber et al., 2017; Madsen et al., 2018; Kleingesse & Gutmann, 2021), which allow a close examination of Bayesian posteriors to identify potential problems or anomalies in the solution.

As we have said, a key problem for model selection is the selection of the *design variables* that define an experiment. When resources are limited, experimental designs can be carefully selected to yield as much information about the models as possible. *Adaptive design optimisation* (ADO) (Cavagnaro et al., 2010, 2013) is one influential approach to selecting experimental designs. ADO proposes designs by maximising the so-called utility objective, which measures the amount of information about the candidate models and their quality. While it is indeed possible for modern methods to approximate common utility objectives, such as mutual information (Cavagnaro et al., 2010; Shannon, 1948) or expected entropy (Yang & Qiu, 2005), it can be challenging when computational models lack a tractable likelihood. In such cases, research suggests adopting LFI methods, in which the computational model generates synthetic observations for inference (Gutmann & Corander, 2016; Sisson et al., 2018; Papamakarios et al., 2019). This broad family of methods is also known as *approximate Bayesian computation* (ABC) (Beaumont et al., 2002; Kangasrääsiö et al., 2019) and simulator- or simulation-based inference (Cranmer et al., 2020). To date, LFI methods for ADO have focused on parameter inference for a single model rather than model selection.

Model selection with limited design iterations requires a choice of design variables that optimise model discrimination as well as improve parameter estimation. The complexity of this task is compounded in the context of LFI, where expensive samples from the model are required. We aim to reduce the number of model simulations. For this reason, in our

approach, called BOSMOS, we use *Bayesian optimisation* (BO) (Frazier, 2018; Greenhill et al., 2020) for both design selection and model selection. The advantage of BO is that it is highly sample-efficient and therefore has a direct impact on reducing the need for model simulation. BOSMOS combines the ADO approach with LFI techniques in a novel way, resulting in a faster method to carry out optimal designs of experiments to discriminate between computational cognitive models with a minimal number of trials.

The main contributions of the paper are as follows:

- A novel approach to simulator-based model selection that casts LFI for multiple models under the Bayesian framework through the approximation of the model likelihood. As a result, the approach provides a full joint Bayesian posterior for models and their parameters given the collected experimental data.
- A novel simulator-based utility objective for choosing experimental designs that maximises the behavioural variation in current beliefs about model configurations. As a part of the adaptive setting, designs are chosen sequentially, each informed by the participant's most recent response. Along with the sample-efficient LFI procedure, the utility objective reduces the time cost from 1 h for competitor methods to less than a minute in the majority of case studies, bringing the method closer to enabling real-time cognitive model testing with human subjects.
- By close integration of the two above contributions, we put forth what we believe to be the first fully Bayesian experimental design approach to model selection that concurrently combines online, sample-efficient, and simulation-based characteristics together in a single, unified methodology.
- The new approach was tested on three well-known paradigms in psychology—memory retention, sequential signal detection, and risky choice—and, despite not requiring likelihoods, reaches similar accuracy to the existing methods that do require them.

Background

In this article, we are concerned with situations where the purpose of experiments is to gather data that can discriminate between models. The traditional approach in such a context begins with the collection of large amounts of data from a large number of participants on a design that is fixed based on intuition; this is followed by evaluation of the model fit using a desired model selection criteria such as the *Akaike information criterion*, the *Bayesian information criterion*, *cross-validation*, etc. This is an inefficient approach—the informativeness of the collected data for choosing models

is unknown in advance, and collecting large amounts of data may often prove expensive in terms of time and monetary resources (for instance, in cases that involve expensive equipment, such as functional magnetic resonance imaging, or in clinical settings). These issues have been addressed by modern optimal experimental design methods, which we consider in this section and summarise in Table 1.

Optimal experimental design OED is a classic problem in statistics (Lindley, 1956; Kiefer, 1959), which saw a resurgence in the last decade due to improvements in computational methods and the availability of computational resources. Specifically, ADO (Cavagnaro et al., 2010, 2013) was proposed for cognitive science models, which have been successfully applied in different experimental settings, including memory and decision-making. In ADO, the designs are selected according to a global utility objective, which is an average value of the local utility over all possible data (behavioural responses) and model parameters, weighted by the likelihood and priors (Myung et al., 2013). More general approaches, such as Kim et al. (2014), improve upon ADO by combining it with hierarchical modelling, which allows them to form richer priors over the model parameters. While useful, the main drawback of these methods is that they work only with tractable (or analytical) parametric models, that is, models whose likelihood is explicitly available and whose evaluation is feasible.

Model selection for simulator-based models In the LFI setting, a critical feature of many cognitive models is that they lack a closed-form solution, but allow forward simulations for a given set of model parameters. A few approaches have made advances in tackling the problem of the intractability of these models. For instance, Kleinegesse and Gutmann

(2020) and Valentin et al. (2021) proposed a method that combines *Bayesian OED* (BOED) and approximate inference of simulator-based models. The *mutual information neural estimation for Bayesian experimental design* (MINEBED) method performs BOED by maximising a lower bound on the expected information gain for a particular experimental design, which is estimated by training a neural network on synthetic data generated by the computational model. By estimating mutual information, the trained neural network no longer needs to model the likelihood directly for selecting designs and doing the Bayesian update. Similarly, *mixed neural likelihood estimation* by Boelts et al. (2022) trains neural density estimators on model simulations to emulate the simulator. Pudlo et al. (2016) proposed an LFI approach to model selection, which uses random forests to approximate the marginal likelihood of the models. Despite these advances, these methods have not been designed for model selection in an adaptive experimental design setting. Table 1 summarises the main differences between modern approaches and the method proposed in this paper.

An alternative way of expressing cognitive models is through an agent-based paradigm (Madsen et al., 2019), where the model can be conceptualised as a *reinforcement learning* (RL) policy (Kaelbling et al., 1996; Sutton & Barto, 2018). The main problem with these agent-based models is that they need retraining if any of their parameters are altered, which introduces a prohibitive computational overhead when doing model selection. Recently, (Moon et al., 2022) proposed a generalised model parameterized by cognitive parameters that can quickly adapt to multiple behaviours, theoretically bypassing the need for model selection altogether and replacing it with parameter inference. Although the cost of evaluating these models is low in general, they

Table 1 A comparison of representative methods for experimental design, with a focus on *parameter estimation* (PE) and *model selection* (MS)

Reference	Method	PE	MS	SBM	Amor	LTP	Adapt	LFI
Cavagnaro et al. (2010)	ADO	✓	✓	x	x	x	✓	x
Kleinegesse and Gutmann (2020)	MINEBED	✓	x	✓	✓	x	x	✓
Blau et al. (2022)	RL-BOED	✓	x	✓	✓	✓	✓	x
Moon et al. (2022)	BOLFI	✓	x	x	x	-	-	✓
Pudlo et al. (2016)	RF-ABC	x	✓	x	✓	-	-	✓
This work	BOSMOS	✓	✓	✓	x	x	✓	✓

Here, we emphasise methods capable of proposing *designs for simulator-based models* (SBM) without the need for tractable model likelihoods. Notably, BOSMOS stands out for its ability to simultaneously handle both PE and MS tasks. Although some methods, such as MINEBED, can recast MS into a PE problem, where the model index is the only parameter and the rest are fixed, in this table, we specifically consider MS for simulator models that also permit changes to their parameters. Key attributes of the presented methods include *amortisation* (Amor.) capabilities, which enable rapid design proposals or inference but require retraining when beliefs update, as in our sequential setting; *long-term planning* (LTP) in design selection is crucial for efficient design space exploration when the budget is known beforehand, as opposed to myopic methods; *adaptive* (Adapt.) design methods accommodate sequential adaptive settings without additional retraining, unlike their static counterparts; and *likelihood-free inference* (LFI) is considered crucial, as we are interested not just in selecting designs for SBM but also in inference. The ‘-’ symbol signifies that design selection is not applicable to the method

lack the interpretability necessary for cognitive theory development. Therefore, training a parameterized policy within a single RL model family may be preferable; this would still require model selection but would avoid the need for retraining when parameters change (see Sect. 4.4 for a concrete example).

Amortised approaches to the OED Recently proposed amortised approaches to OED (Blau et al., 2022)—i.e., flexible machine learning models trained upfront on a large set of problems with the goal of making fast design selection at runtime—allow more efficient selection of experimental designs by introducing an RL policy that generates design proposals. This policy provides a better exploration of the design space, does not require access to a differentiable probabilistic model, and can handle both continuous and discrete design spaces, unlike previous amortised approaches (Foster et al., 2021; Ivanova et al., 2021). These amortised methods have yet to be applied to model selection.

Even though OED is a classical problem in statistics, its application has mostly been relegated to discriminating between simple, tractable models. Modern methods such as LFI and amortised inference can, however, make it more feasible to develop OED methods that can work with complex simulator models. In the next sections, we elaborate on our LFI-based method BOSMOS and demonstrate its working using three classical cognitive science tasks: memory retention, sequential signal detection, and risky choice.

Methods

Our method carries out optimal experiment design for model selection and parameter estimation, involving three main stages as shown in Fig. 1: selecting the experimental design \mathbf{d} , collecting new data \mathbf{x} at the design \mathbf{d} chosen from a design space, and, finally, updating current beliefs about the models and their parameters. The process continues until the allocated budget for design iterations T is exhausted and the preferred cognitive model $m_{\text{est}} \in \mathcal{M}$, which explains the subject behaviour the best, and its parameters $\theta_{\text{est}} \in \Theta_{\text{est}}$ are extracted. While the method is rooted in Bayesian inference and thus builds a full joint posterior over models and parameters, we also consider that ultimately the experimenter may want to report the single *best* model and parameter setting, and we use this decision-making objective to guide the choices of our algorithm. The definition of what ‘best’ here means depends on a cost function chosen by the user (Robert, 2007). In this paper, for the sake of simplicity, we choose the most common Bayesian estimator, the *maximum a posteriori*

(MAP), of the *full posterior* computed by the method:

$$m_{\text{est}} = \arg \max_m p(m \mid \mathcal{D}_{1:t}), \quad (1)$$

$$\theta_{\text{est}} = \arg \max_{\theta_m} p(\theta_m \mid m, \mathcal{D}_{1:t}), \quad (2)$$

where $m \in \mathcal{M}$, $\theta_m \in \Theta_m$ and $\mathcal{D}_{1:t} = ((\mathbf{d}_1, \mathbf{x}_1), \dots, (\mathbf{d}_t, \mathbf{x}_t))$ is a sequence of experimental designs \mathbf{d} (e.g., shown stimulus) and the corresponding behavioural data \mathbf{x} (e.g., the response of the subject to the stimuli) pairs. It is noteworthy that while we illustrate the BOSMOS methodology using the MAP rule, it is flexible to accommodate alternative estimators if required.

Key assumptions In our usage context, it is important to make a few reasonable assumptions. First, we assume that the prior over the models $p(m)$ and their parameters $p(\theta_m \mid m)$, as well as the domain of the design space, have been specified using sufficient prior knowledge; they may be given by expert psychologists or previous empirical work. This guarantees that the space of the problem is well defined. Notice that this also implies that the set of candidate models $\mathcal{M} = (m_1, \dots, m_k)$ is known, and each model is defined, for any design, by its own parameters. Second, we assume that the computational models that we consider may not necessarily have a closed-form solution, in case their likelihoods $p(\mathbf{x} \mid \mathbf{d}, \theta_m, m)$ are intractable but it is possible to sample from the forward model m given the parameter setting θ_m , and design \mathbf{d} . In other words, we operate in a simulator-based inference setting. Please note that this likelihood depends only on the current design and parameters, as assumed in our setting; however, in general, OED techniques (including BOSMOS) can handle scenarios where experimental trials are not strictly independent, provided the participant’s behaviour changes are explicitly accommodated by a behaviour model. The third assumption is that each subject’s dataset is analysed separately: we consider single subjects with fixed parameters undergoing the whole set of experiments, as opposed to the statistical setting where information about one dataset may impact the whole population, such as, for instance, in hierarchical modelling or pooled models.

Sequential design selection and belief updates As evidenced by Eqs. 1 and 2, the sequential choice of the designs at any point depends on the current posterior over the models and parameters $p(\theta_m, m \mid \mathcal{D}_{1:t}) = p(\theta_m \mid \mathcal{D}_{1:t}, m) \cdot p(m \mid \mathcal{D}_{1:t})$, which needs to be approximated and updated at each iteration step of the main loop in Fig. 1. This problem can be formulated through sequential importance sampling methods, such as *sequential Monte Carlo* (SMC; (Del Moral et al., 2006)).

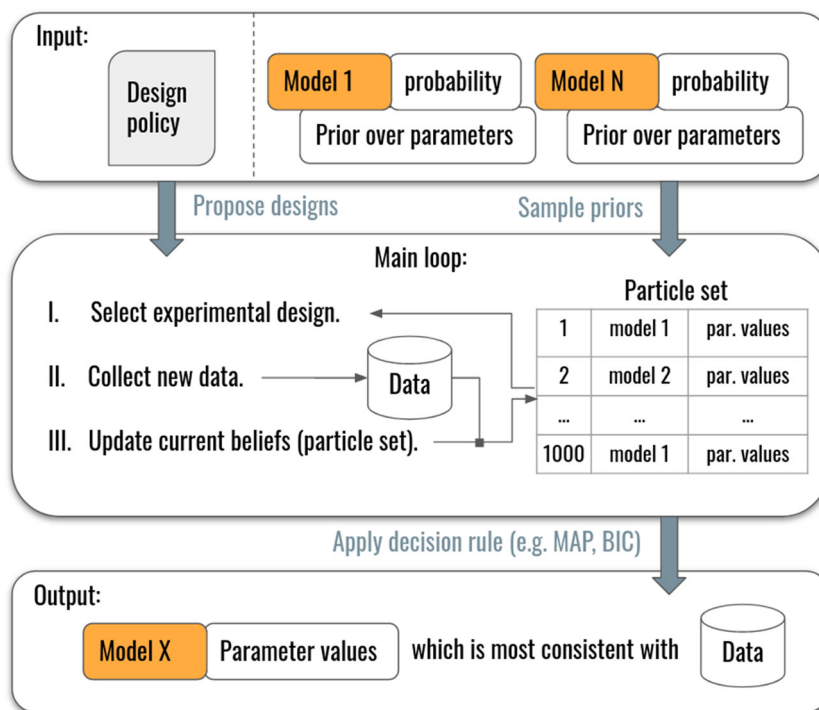


Fig. 1 Components of the model selection approach. The main loop continues until the experimental design budget is depleted. Input panel: the experimenter defines a design policy (e.g., random choice of designs), as well as the models and their parameter priors. Middle panel: (i) the next experimental design is selected based on the design policy and current beliefs about models and their parameters (initially sampled from model and parameter priors); (ii) the experiment is carried

out using the chosen design, and the observed response-design pair is stored; (iii) current beliefs are updated (e.g., resampled) based on experimental evidence acquired thus far. Output panel: the model and parameters that are most consistent with the collected data are selected by applying one of the well-established decision rules to the final beliefs about models and their parameters

Thus, the resulting parameter posteriors can be approximated, up to resampling, in the form of equally weighted particle sets: $q_t(\theta_m, m \mid \mathcal{D}_{1:t}) = \sum_{i=1}^{N_1} N_1^{-1} \delta_{\theta_m^{(i)}, m^i}$, with $\theta_m^{(i)}, m^{(i)}$ the parameters and models associated with the particle i , as an approximation of $p(\theta_m \mid m, \mathcal{D}_{1:t})$. These particle sets are later sampled to select designs and update parameter posteriors.

Preventing particle degeneracy An important consideration when using this method, similar to all particle methods, is the potential for particle degeneracy, a situation where a few particles disproportionately represent the posterior (Doucet et al., 2000; Liu & Chen, 1998). To mitigate these issues, several strategies can be applied, including regular checks of the *effective sample size* (ESS), a diagnostic measure of particle diversity, with lower values indicating potential degeneracy (Kong et al., 1994). When ESS falls below a specified threshold, resampling procedures such as systematic or stratified resampling can be initiated, which redistribute weights across particles to prevent degeneracy (Doucet et al., 2000). Additional regularisation techniques can be employed too, introducing small amounts of artificial noise to the particle to prevent particle dominance (Musso et al., 2001), and multiple

proposal distributions can be used to offer varied *pathways* for particle movements, reducing the risk of particle clustering (Doucet et al., 2000). In this paper, we use resampling and regularisation described in Sect. 4.1, and we further detail this problem of particle degeneracy in Appendix D. In the following sections, we take a closer look at the design selection and belief update stages.

Selecting Experimental Designs

Traditionally, in the experimental design literature, the designs are selected at each iteration t by maximising the reduction of the expected entropy $H(\cdot)$ of the posterior $p(m, \theta_m \mid \mathcal{D}_{1:t})$. By definition of conditional probability, we have the following:

$$d_t = \operatorname{argmin}_{d_t} \mathbb{E}_{x_t \mid \mathcal{D}_{1:t-1}} [H(\theta_m, m \mid \mathcal{D}_{1:t-1} \cup (d_t, x_t))] \tag{3}$$

$$= \operatorname{argmin}_{d_t} \mathbb{E}_{x_t \mid \mathcal{D}_{1:t-1}} [\mathbb{E}_{p(\theta_m, m \mid \mathcal{D}_{1:t})} [-\log p(\theta_m, m \mid \mathcal{D}_{1:t} \cup (d_t, x_t))]]$$

$$= \operatorname{argmin}_{d_t} \mathbb{E}_{x_t \mid \mathcal{D}_{1:t-1}} \mathbb{E}_{p(\theta_m, m \mid \mathcal{D}_{1:t-1})} [-\log(p(x_t \mid d_t, \theta_m, m))] + \mathbb{E}_{x_t \mid \mathcal{D}_{1:t-1}} \log p(x_t \mid d_t, \mathcal{D}_{1:t-1}), \tag{4}$$

where x_t is the response predicted by the model. The first equality comes from the definition of entropy, and the second

from Bayes rule, where we removed the prior, as this term is a constant term in \mathbf{d}_t . Here, lower entropy corresponds to a narrower, more concentrated posterior—with maximal information about models and parameters.

Since neither $p(\mathbf{x}_t | \mathbf{d}_t, \boldsymbol{\theta}_m, m)$ nor, by extension, Eq. 4 are tractable in our setting, we propose a simulator-based utility objective:

$$\mathbf{d}_t = \operatorname{argmin}_{\mathbf{d}_t} \mathbb{E}_{q_t(\boldsymbol{\theta}_m, m | \mathcal{D}_{1:t-1})} [\hat{H}(\mathbf{x}'_t | \mathbf{d}_t, \boldsymbol{\theta}_m, m)] - \hat{H}(\mathbf{x}_t | \mathcal{D}_{1:t-1}, \mathbf{d}_t), \quad (5)$$

where q_t is a particle approximation of the posterior at time t , and \hat{H} is a kernel-based Monte Carlo approximation of the entropy H .

The intuition behind this utility objective is that we choose such designs \mathbf{d}_t that would maximise identifiability (minimise the entropy) between N responses \mathbf{x}' simulated from different computational models $p(\cdot | \mathbf{d}_t, \boldsymbol{\theta}_m, m)$. The models m as well as their parameters $\boldsymbol{\theta}_m$ are sampled from the current beliefs $q_t(\boldsymbol{\theta}_m, m | \mathcal{D}_{1:t-1})$. This utility objective balances model and parameter exploration through design choices guided by the posterior distribution. As the method continuously assesses the fluctuating uncertainty levels in each space, it adapts and reassigns priorities to maintain an efficient exploration strategy, addressing a potential conflict between the two objectives of model selection and parameter estimation. The full asymptotic validity of the Monte Carlo approximation of the decision rule in Eq. 5 can be found in Appendix A.

The utility objective in Eq. 5 allows us to use BO to find the design \mathbf{d}_t and then run the experiment with the selected design. In the next section, we discuss how to update beliefs about the models m and their parameters $\boldsymbol{\theta}_m$ based on the data collected from the experiment.

Likelihood-Free Posterior Updates

The response \mathbf{x}_t from the experiment with the design \mathbf{d}_t is used to update approximations of the posterior $q_t(m | \mathcal{D}_t)$ and $q_t(\boldsymbol{\theta}_m | m, \mathcal{D}_t)$, obtained via marginalisation and conditioning, respectively, from $q_t(\boldsymbol{\theta}_m, m | \mathcal{D}_t)$. We use LFI with synthetic responses $\mathbf{x}_{\boldsymbol{\theta}_m}$ simulated by the behavioural model m to perform the approximate Bayesian update.

Parameter estimation conditioned on the model We start with parameter estimation for each of the candidate models using *BO for LFI* (BOLFI; (Gutmann & Corander, 2016)). In BOLFI, a *Gaussian process* (GP) (Rasmussen, 2004) surrogate for the discrepancy function between the observed and simulated data, $\rho(\mathbf{x}_{\boldsymbol{\theta}_m}, \mathbf{x}_t)$ (e.g., Euclidean distance), serves as a base to an unnormalized approximation of the intractable likelihood $p(\mathbf{x}_t | \mathbf{d}_t, \boldsymbol{\theta}_m, m)$. Thus, the posterior

can be approximated through the following approximation of the likelihood function $\mathcal{L}_{\epsilon_m}(\cdot)$ and the prior over model parameters $p(\boldsymbol{\theta}_m)$:

$$p(\boldsymbol{\theta}_m | \mathbf{x}_t) \propto \mathcal{L}_{\epsilon_m}(\mathbf{x}_t | \boldsymbol{\theta}_m) \cdot p(\boldsymbol{\theta}_m), \quad (6)$$

$$\mathcal{L}_{\epsilon_m}(\mathbf{x}_t | \boldsymbol{\theta}_m) \approx \mathbb{E}_{\mathbf{x}_{\boldsymbol{\theta}_m}} [\kappa_{\epsilon_m}(\rho_m(\mathbf{x}_{\boldsymbol{\theta}_m}, \mathbf{x}_t))]. \quad (7)$$

Here, following Section 6.3 of Gutmann and Corander (2016), we choose $\kappa_{\epsilon_m}(\cdot) = \mathbf{1}_{[0, \epsilon_m]}(\cdot)$, where the bandwidth ϵ_m takes the role of an acceptance-rejection threshold. Using a Gaussian likelihood for the GP, this leads to $\mathbb{E}_{\mathbf{x}_{\boldsymbol{\theta}_m}} [\kappa_{\epsilon_m}(\rho(\mathbf{x}_{\boldsymbol{\theta}_m}, \mathbf{x}_t))] = \Phi((\epsilon_m - \mu(\boldsymbol{\theta}_m)) / \sqrt{v(\boldsymbol{\theta}_m) + \sigma^2})$, where $\Phi(\cdot)$ denotes the standard Gaussian *cumulative distribution function* (CDF). Note that $\mu(\boldsymbol{\theta}_m)$ and $v(\boldsymbol{\theta}_m) + \sigma^2$ are the posterior predictive mean and variance of the GP surrogate at $\boldsymbol{\theta}_m$.

Model estimation A principled way of performing model selection is via the marginal likelihood, that is $p(\mathbf{x}_t | m) = \int p(\mathbf{x}_t | \boldsymbol{\theta}_m, m) \cdot p(\boldsymbol{\theta}_m | m) d\boldsymbol{\theta}_m$, which is proportional to the posterior over models assuming an equal prior for each model. Unfortunately, a direct computation of the marginal likelihood is not possible with Eq. 7, since it only allows us to compute a likelihood approximation up to a scaling factor that implicitly depends on ϵ . For instance, when calculating a Bayes factor (ratio of marginal likelihoods) for models m_1 and m_2 ,

$$\frac{p(\mathbf{x}_t | m_1)}{p(\mathbf{x}_t | m_2)} = \frac{\mathbb{E}_{\boldsymbol{\theta}_{m_1}} [p(\mathbf{x}_t | \boldsymbol{\theta}_{m_1}, m_1)]}{\mathbb{E}_{\boldsymbol{\theta}_{m_2}} [p(\mathbf{x}_t | \boldsymbol{\theta}_{m_2}, m_2)]} \neq \frac{\mathbb{E}_{\boldsymbol{\theta}_{m_1}} [\mathcal{L}_{\epsilon_{m_1}}(\mathbf{x}_t | \boldsymbol{\theta}_{m_1})]}{\mathbb{E}_{\boldsymbol{\theta}_{m_2}} [\mathcal{L}_{\epsilon_{m_2}}(\mathbf{x}_t | \boldsymbol{\theta}_{m_2})]}, \quad (8)$$

their respective ϵ_{m_1} and ϵ_{m_2} , chosen independently, may potentially bias the marginal likelihood ratio in favour of one of the models, rendering it unsuitable for model selection. Choosing the same ϵ for each model is not possible either, as it would lead to numerical instability due to the shape of the kernel.

To approximate the marginal likelihood $p(\mathbf{x}_t | m)$, we adopt a similar approach as in Eq. 7, by reframing the marginal likelihood computation as a distinct LFI problem. In ABC, for parameter estimation, we would generate pseudo-observations from the prior predictive distribution of each model and compare the discrepancy with the true observations on a scale common to all models. This comparison involves a kernel that maps the discrepancy into a likelihood approximation. For example, in rejection ABC (Tavaré et al., 1997; Marin et al., 2012) this kernel is uniform. In our case, we will generate samples from the joint prior predictive distribution on both models and parameters, and we use a Gaussian kernel $\kappa_\eta(\cdot) = \mathcal{N}(\cdot | 0, \eta^2)$, chosen to satisfy all of the requirements from Gutmann & Corander (2016); in particular, this kernel is non-negative,

non-concave, and has a maximum at 0. The parameter $\eta > 0$ serves as the kernel bandwidth, similarly to ϵ_m in Eq. 7. The value of $\kappa_\eta(\cdot)$ monotonically increases as the model m produces smaller discrepancy values. This kernel leads to the following approximation of the marginal likelihood:

$$\mathcal{L}(\mathbf{x}_t \mid m, \mathcal{D}_{t-1}) \propto \mathbb{E}_{\mathbf{x}_\theta \sim p(\cdot \mid \theta_m, m) \cdot q(\theta_m \mid m, \mathcal{D}_{t-1})} \kappa_\eta(\hat{\rho}(\mathbf{x}_\theta, \mathbf{x}_t)), \tag{9}$$

where $\kappa_\eta(\cdot) = \mathcal{N}(\cdot \mid 0, \eta^2)$, and $\hat{\rho}$ is the GP surrogate for the discrepancy. Equation 9 is a direct equivalent of Eq. 7, but here we integrate (marginalise) over both θ and x_θ . Here we used the Gaussian kernel instead of the uniform kernel used in Eq. 7, as it produced better results for model selection in preliminary numerical experiments. Note that in Eq. 9 we have two approximations, the first one from κ_η , stating that the likelihood is approximated from the discrepancy, and the second from the use of a GP surrogate for the discrepancy.

The choice of η is a complex problem, and in this paper we propose the simple solution of setting η as the minimum value of $\mathbb{E}_{\mathbf{x}_\theta \sim p(\cdot \mid \theta_m, m) \cdot q(\theta_m \mid m, \mathcal{D}_{t-1})} \hat{\rho}(\mathbf{x}_\theta, \mathbf{x}_t)$ across all models $m \in \mathcal{M}$. This value has the advantage of giving non extreme values to the estimations of the marginal likelihood, which should in principle avoid overconfidence.

Posterior update The resulting marginal likelihood approximation in Eq. 9 can then be used in posterior updates for new design trials as follows:

$$q(m \mid \mathcal{D}_t) \propto \mathcal{L}(\mathbf{x}_t \mid m, \mathcal{D}_{t-1}) \cdot q(m \mid \mathcal{D}_{t-1}) \approx \kappa_\eta(\omega_m) \cdot q(m \mid \mathcal{D}_{t-1}), \tag{10}$$

$$q(\theta_m \mid m, \mathcal{D}_t) \propto \mathcal{L}_{\epsilon_m}(\mathbf{x}_t \mid \theta_m, m) \cdot q(\theta_m \mid \mathcal{D}_{t-1}, m), \tag{11}$$

which is equivalent to the following:

$$q(\theta_m, m \mid \mathcal{D}_t) \propto \mathcal{L}_{\epsilon_m}(\mathbf{x}_t \mid \theta_m, m) \cdot \mathcal{L}(\mathbf{x}_t \mid m, \mathcal{D}_{t-1}) \cdot q(\theta_m, m \mid \mathcal{D}_{t-1}). \tag{12}$$

Once we update the joint posterior of models and parameters, it is straightforward to obtain the model and parameter posterior through marginalisation and apply a decision rule (e.g., MAP) to choose the estimate. The entire algorithm for BOSMOS can be found in Appendix B.

Experiments

In the experiments, our goal was to evaluate how well the proposed method described in Sect. 3 discriminated between different computational models in a series of cognitive tasks: memory retention, signal detection, and risky choice. Specifically, we measured how well the method chooses designs that help the estimated model imitate the behaviour of the target model, discriminate between models, and correctly estimate their ground-truth parameters. In our simulated experimental setup, we created 100 synthetic participants by sampling the

ground-truth model and its parameters (not available in the real world) through priors $p(m)$ and $p(\theta_m \mid m)$. Then, we ran the sequential experimental design procedure for a range of methods described in Sect. 4.1, and recorded four main performance metrics shown in Fig. 3 for 20 design trials (results analysed further later in the section): the behavioural fitness error η_b , defined below, the parameter estimation error η_p , the accuracy of the model prediction η_m and the empirical time cost of running the methods. Furthermore, we evaluated the methods at different stages of design iterations in Fig. 3 for the convergence analysis. The complete experiments, with additional evaluation points and details about hardware, can be found in Appendix C.

We compute η_b , η_p and η_m for a single synthetic participant using the known ground truth model m_{true} and parameters θ_{true} . The behavioural fitness error $\eta_b = \|\mathbf{X}_{\text{true}} - \mathbf{X}_{\text{est}}\|^2$ is calculated as the Euclidean distance between the ground-truth model (\mathbf{X}_{true}) and synthetic (\mathbf{X}_{est}) behavioural datasets, which consist of means $\mu(\cdot)$ of 100 responses evaluated at the same 100 random designs \mathcal{T} generated from a proposal distributions $p(\mathbf{d})$, defined for each model:

$$\mathcal{T} = \{\mathbf{d}_i \sim p(\mathbf{d})\}_{i=1}^{100}, \tag{13}$$

$$\mathbf{X}_{\text{true}} = \{\mu(\{\mathbf{x}_s : \mathbf{x} \sim p(\cdot \mid \mathbf{d}_i, \theta_{\text{true}}, m_{\text{true}})\}_{s=1}^{100}) : \mathbf{d}_i \in \mathcal{T}\}_{i=1}^{100}, \tag{14}$$

$$\mathbf{X}_{\text{est}} = \{\mu(\{\mathbf{x}_s : \mathbf{x} \sim p(\cdot \mid \mathbf{d}_i, \theta_{\text{est}}, m_{\text{est}})\}_{s=1}^{100}) : \mathbf{d}_i \in \mathcal{T}\}_{i=1}^{100}. \tag{15}$$

Here, m_{est} and θ_{est} are, respectively, the model and parameter values estimated via the MAP rule (unless specified otherwise). m_{est} is also used to calculate the predictive model accuracy η_m as a proportion of correct model predictions for the total number of synthetic participants, while θ_{est} is used to calculate the averaged Euclidean distance $\|\theta_{\text{true}} - \theta_{\text{est}}\|^2$ across all synthetic participants, which constitutes the parameter estimation error η_p .

Comparison Methods

Throughout the experiments, we compare several strategies for experimental design selection and parameter inference. A prior predictive distribution with a random design choice drawn from each model’s proposal distribution serves as our baseline, referred to as the *Prior* in the results.

As we implemented these strategies, one of the key focuses was mitigating the risk of particle degeneracy. In all our particle methods, we have included safety measures such as resampling the beliefs represented by the particles, denoted as $q_t(\theta_m, m \mid \mathcal{D}_{1:t})$, after each design trial. Additionally, we have incorporated a regularisation technique that introduces a

small amount of artificial Gaussian noise to the particles during resampling. These measures enhance the robustness of methods against particle degeneracy, ensuring a more reliable and stable inference. The precise details of methods, along with the specific setup parameters, are elaborated below.

Likelihood-Based Inference with Random Design

Likelihood-based inference with random design (LBIRD) applies the ground-truth likelihood, where it is possible, to conduct Bayesian inference and samples the design from the proposal distribution $p(\mathbf{d})$ instead of design selection: $\mathcal{D}_t = (x_t, \mathbf{d}_t)$, $x_t \sim \pi(\cdot | \boldsymbol{\theta}, m, \mathbf{d}_t)$, $\mathbf{d}_t \sim p(\cdot)$. This procedure serves as a baseline by providing unbiased estimates of models and parameters. As other methods in this section, LBIRD uses 5000 particles (empirical samples) to approximate the joint posterior of models and parameters for each model. The Bayesian updates are conducted through *importance-weighted* sampling with added Gaussian noise applied to the current belief distribution.

ADO

ADO requires a tractable likelihood of the models and is hence used as an upper bound of performance in cases where the likelihood is available. ADO (Cavagnaro et al., 2010) employs BO for the mutual information utility objective:

$$U(\mathbf{d}) = \sum_{m=1}^K p(m) \sum_y p(x | m, \mathbf{d}) \cdot \log \left(\frac{p(x | m, \mathbf{d})}{\sum_{m=1}^K p(m) p(x | m, \mathbf{d})} \right), \quad (16)$$

where we used 500 parameters sampled from the current beliefs to integrate

$$p(x | m, \mathbf{d}) = \int p(x | \boldsymbol{\theta}_m, m, \mathbf{d}) \cdot p(\boldsymbol{\theta}_m | m) d\boldsymbol{\theta}. \quad (17)$$

Similarly to other approaches below, which also use BO, the BO procedure is initialised with 10 evaluations of the utility objective with \mathbf{d} sampled from the design proposal distribution $p(\mathbf{d})$, while the next 5 design locations are determined by the *Monte-Carlo-based noisy expected improvement* objective. The GP surrogate for the utility uses a constant mean function, a Gaussian likelihood, and the *Matern* kernel with zero mean and unit variance. All these components of the design selection procedure were implemented using the *BOTorch* package (Balandat et al., 2020).

MINEBED

MINEBED, as presented by Kleinegesse and Gutmann (2020), is a technique that specialises in design selection for parameter inference within a single model framework or

for model selection among models with a predetermined set of parameter values. This stands in contrast to our requirement for both model selection and parameter inference at the same time and, by extension, working with multiple models where the parameter values are not fixed but rather allowed to vary and be estimated. To accommodate this discrepancy, we operate separate MINEBED instances for each model and delegate design optimisation to a single model-chosen from current beliefs-at each trial. The model is selected using the MAP rule over the current beliefs about models $q(m | \mathcal{D}_{1:t})$. The data obtained from conducting the experiment with the selected design are then used to update all MINEBED instances.

Given that MINEBED was originally developed for static experimental designs, to fit it into the adaptive setting of our experiments, we adopted a strategy of retraining each MINEBED instance from scratch with updated beliefs after each design trial to propose the next design. This adaptation, while necessary for our use case, does introduce a notable difference from the standard MINEBED approach. In particular, it significantly increases the computation cost of applying the method.

The specific MINEBED implementation we employed is based on the original work by Kleinegesse and Gutmann (2020), utilising a neural surrogate for mutual information made up of two fully connected layers with 64 neurons each. This surrogate was optimised using the *Adam* optimizer (Kingma & Ba, 2014) with an initial learning rate of 0.001, 5000 simulations per training at each new design trial, and 5000 epochs.

It is worth mentioning that MINEBED, as an instance of the broader method of using mutual information estimation for Bayesian experimental design (Michaud, 2019), is not strictly tied to a single configuration. In the original paper, the authors perform a small scale optimisation of the hyperparameters (in particular learning rate and depth of the neural network), leading to similar optimal values in the main examples. However, the settings of our study, involving 1–4 parameters and 1–4 designs, are comparable to the original study's conditions in Kleinegesse and Gutmann (2020), which handled 2–4 parameters and a single design. Thus, we chose a value close to the ones proposed in the original paper.

BOSMOS

BOSMOS is the method proposed in this paper and described in Sect. 3. It uses the simulator-based utility objective from Eq. 5 in BO to select the design and BO for LFI, along with the marginal likelihood approximation from Eq. 9 to conduct inference. The objective for design selection is calculated with the same 10 models (a higher number increases belief representation at the cost of more computations) sampled from the current belief over models (i.e., particle set $q_t(m |$

$\mathcal{D}_{1:t}$) at each time t), where each model is simulated 10 times to get one evaluation point of the utility (100 simulations per point). In total, in each iteration, we spent 1500 simulations to select the design and an additional 100 simulations to conduct parameter inference.

As for parameter inference in BOSMOS, BO was initialised with 50 parameter points randomly sampled from the current beliefs about model parameters (i.e., the particle set $q_t(\theta_m | m, \mathcal{D}_{1:t})$), the other 50 points were selected for simulation in batches of 5 through the *lower confidence bound selection criteria* (Srinivas et al., 2009) acquisition function. Once again, a GP is used as a surrogate, with the constant mean function and the *radial basis function* (Seeger, 2004) kernel with zero mean and unit variance. Once the simulation budget of 100 is exhausted, the parameter posterior is extracted through an *importance-weighted* sampling procedure, where the GP surrogate with the tolerance threshold set at a minimum of the GP mean function (Gutmann & Corander, 2016) acts as a base for the simulator parameter likelihood.

Demonstrative Example

The demonstrative example serves to highlight the significance of design optimisation for model selection with a simple toy scenario. This example incorporates two normal distribution models: the *positive mean* (PM) model and the *negative mean* (NM) model.

The PM and NM models generate responses influenced by the experimental design d , which determines the amount of observational noise variance. The proposal distribution for d is defined as $d \sim \text{Unif}(0.001, 5)$. Under this setting, the PM and NM models are formally described as follows:

$$\text{PM} : x \sim \mathcal{N}(\theta_\mu, d^2); \quad \text{NM} : x \sim \mathcal{N}(-\theta_\mu, d^2). \quad (18)$$

Both models share a common uniform prior over their single parameter θ_μ , which is defined by $\theta_\mu \sim \text{Unif}(0, 5)$. It is worth noting that the models can be distinctly differentiated when the optimal design value is set at $d = 0.001$. Finally, for this example, we employ a uniform prior over the models themselves.

Results

As shown in the first set of analyses in Fig. 2, selecting informative designs can be crucial. When compared to the LBIRD method, which picked designs at random, all the design optimisation approaches performed exceedingly well. This highlights the significance of design selection, as random designs produce uninformative results and impede the inference procedure.

Figure 3 illustrates the convergence of the key performance measures, demonstrating that the design optimisation methods had nearly perfect estimates of ground-truths after only one design trial. This indicates that the PM and NM models are easily separable, provided they have informative designs. In terms of the model predictive accuracy, MINEBED outperformed BOSMOS after the first trial; however, BOSMOS rapidly caught up as trials proceeded. This is most likely because our technique employs fewer simulations per trial but a more efficient LFI surrogate than MINEBED. As a result, our method has the second-best time cost not only for the demonstrative example but also across all cognitive tasks. The only method that was faster is the LBIRD method, which skips the design optimisation procedure entirely and avoids lengthy computations related to LFI by accessing the ground-truth likelihood.

Memory Retention

Studies of memory are a fundamental research area in experimental psychology. Memory can be viewed functionally as a capability to encode, store, and remember and neurologically as a collection of neural connections (Amin & Malik, 2013). Studies of memory retention have a long history in psychological research, in particular in relation to the shape of the retention function (Rubin & Wenzel, 1996). These studies on functional forms of memory retention seek to quantitatively answer how long a learned skill or material is available (Rubin et al., 1999) or how quickly it is forgotten. Distinguishing retention functions may be a challenge (Rubin et al., 1999), and Cavagnaro et al. (2010) showed that employing an ADO approach can be advantageous. Specifically, studies of memory retention typically consist of a *study phase* (for memorising) followed by a *test phase* (for recalling), and the time interval between the two is called a *lag time*. Varying the lag time by means of ADO allowed more efficient differentiation of the candidate models (Cavagnaro et al., 2010). To demonstrate our approach with the classic memory retention task, we consider the case of distinguishing two functional forms, or models, of memory retention, defined as follows.

Models In the classic memory retention task, the subject is tasked with recalling a specific stimulus, such as a word, after a certain time span d . The time variable d serves as a design variable with a proposal distribution $d \sim \text{Unif}(0, 100)$.

The memory processed is modelled using two Bernoulli models: the *power* (POW) model and the *exponential* (EXP) model. The resultant samples from these models, denoted by x , correspond to the responses to the task. An outcome of $x = 0$ implies that the stimulus has been forgotten, while $x = 1$ signifies successful recall.

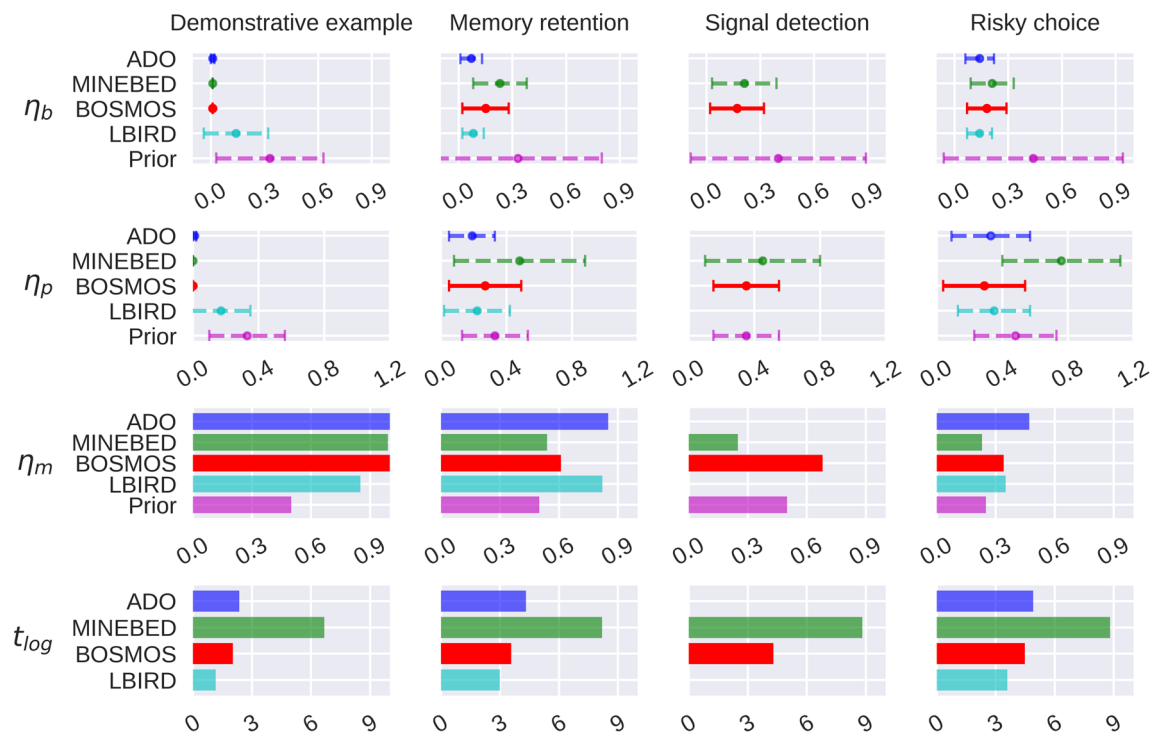


Fig. 2 Comparison of various method performances (rows) after 20 design trials across four cognitive modelling tasks (columns): demonstrative example, memory retention, signal detection, and risky choice. Our proposed BOSMOS method (red) consistently outperforms the alternative LFI technique, MINEBED (green), requiring ten times fewer simulations and significantly reducing time costs by a factor of 60 to 100. Key performance indicators include behavioural fitness error η_b , parameter estimation error η_p , model predictive accuracy η_m , and empirical time cost t_{log} . Notably, these timings, displayed here in minutes on

a logarithmic scale for a total of 100 designs, translate to individual trial intervals of 0.2 to 0.54 min for BOSMOS on the hardware specified in Appendix C.4. Model accuracy bars represent the proportion of correct model predictions across 100 simulated participants, while error bars indicate the mean (marker) and standard deviation (caps) of errors. Note that ABO (dark blue) and LBIRD (cyan) methods are absent in the third column due to their inability to handle models in the sequential signal detection task due to their lack of straightforward likelihood approximations

We follow the definition of these models by Cavagnaro et al. (2010), where a probability p of remembering the stimulus is modelled as follows:

$$x \sim B(1, p), \quad (19)$$

$$\mathbf{POW} : p = \theta_a \cdot (d+1)^{-\theta_{\text{POW}}}; \quad \mathbf{EXP} : p = \theta_a \cdot e^{-\theta_{\text{EXP}} \cdot d}, \quad (20)$$

For both models, the prior probabilities of the parameters are given as:

$$\theta_a \sim \text{Beta}(2, 1), \quad \theta_{\text{POW}} \sim \text{Beta}(1, 4), \quad \theta_{\text{EXP}} \sim \text{Beta}(1, 8). \quad (21)$$

Similarly to the previous demonstrative example and the rest of the experiments, we maintain an equal prior probability distribution across all models.

Results

Studies on the memory task show that the performance gap between LFI approaches and methods that use ground-truth likelihood grows as the number of design trials increases (Fig. 2). This is expected since doing LFI introduces an approximation error, which becomes more difficult to decrease when the most uncertainty around the models and their parameters has already been removed by previous trials. Unlike in the demonstrative example, where design selection was critical, the ground-truth likelihood appears to have a larger influence than design selection for this task, as evidenced by the similar performance of the LBIRD and ADO approaches.

In regard to LFI techniques, BOSMOS outperforms MINEBED in terms of behavioural fitness and parameter estimation, as shown in Fig. 3, but only marginally better for model selection. Moreover, both approaches seem to con-

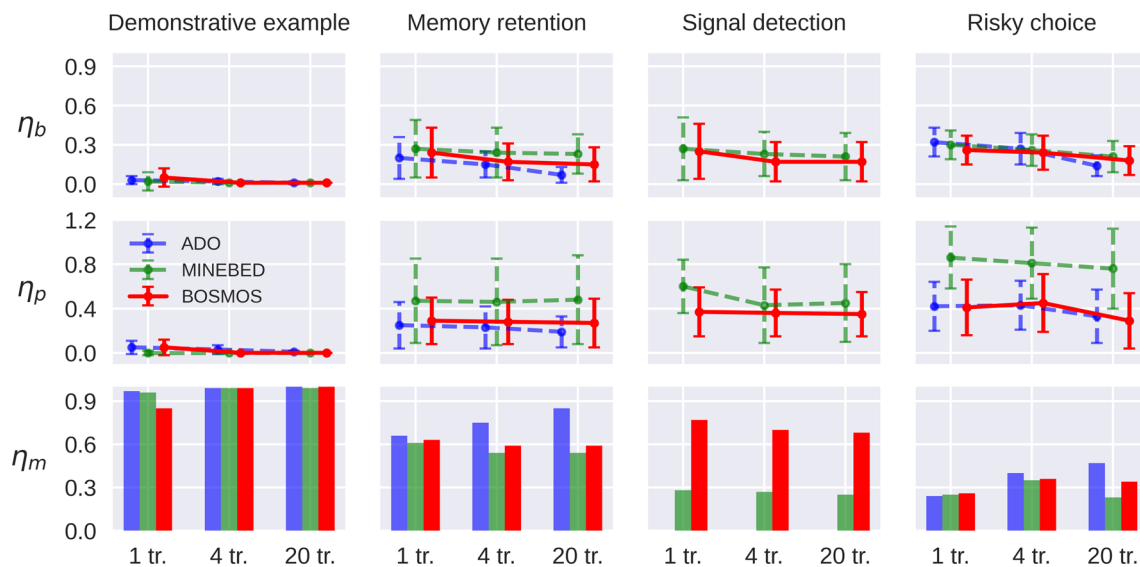


Fig. 3 Evaluation of three performance measures (rows) after 1, 4, and 20 design trials for BOSMOS (solid red) and two alternative best methods, ADO (blue) and MINEBED (green), in four cognitive tasks (columns). As the number of design trials grows, the methods accumulate more observed data from subjects' behaviour and, hence, should reduce behavioural fitness error η_b , parameter estimation error η_p , and increase model predictive accuracy η_m . Since η_b is the per-

formance metric MINBED and BOSMOS optimise, its convergence is the most prominent. The lack of convergence for the other two metrics in the memory retention and signal detection tasks is likely due to the possibility of the same behavioural data being produced by models and parameters that are different from the ground-truth (i.e., non-identifiability of these models)

verge to the wrong solutions (unlike ADO), as evidenced by their lack of convergence in the parameter estimation and model accuracy plots. Interestingly, both techniques continued improving behavioural fitness, implying that the behavioural data of the models can be reproduced by several parameters that are different from the ground-truth, and LFI methods fail to distinguish them. A deeper examination of the parameter posterior can reveal this issue, which can likely be alleviated by adding new features for observations and designs that can assist in capturing the intricacies within the behavioural data.

Sequential Signal Detection

Signal detection theory (SDT) focuses on perceptual uncertainty, presenting a framework for studying decisions under such ambiguity (Tanner & Swets, 1954; Peterson et al., 1954; Swets et al., 1961; Wickens, 2002). SDT is an influential and developing model stemming from mathematical psychology and psychophysics, providing an analytical framework for assessing optimal decision-making in the presence of ambiguous and noisy signals. The origins of SDT can be traced to the 1800s, but its modern form emerged in the latter half of the twentieth century with the realisation that sensory noise is consciously accessible (Wixted, 2020). An example of a signal detection task could be a doctor making a diag-

nosis: they have to make a decision based on a (noisy) signal of different symptoms (Wickens, 2002). Our approach to the sequential signal detection task is rooted in the normative belief that decision-makers operate within rational bounds (Swets et al., 1961). We consider two models in this context: *proximal policy optimisation* (PPO) and *probability ratio* (PR). These models follow the methodology used for computational rational participants, which has been demonstrated to capture a range of behaviours as discussed by Howes et al. (2009).

Task Description

In the signal detection task, the subject needs to correctly discriminate the presence of the signal $o_{\text{sign}} \in \{\text{present}, \text{absent}\}$ in a sensory input $o_{\text{in}} \in \mathbb{R}$. The sensory input is corrupted with sensory noise $\sigma_{\text{sens}} \in \mathbb{R}$:

$$o_{\text{in}} = \mathbb{1}_{\text{present}}(o_{\text{sign}}) \cdot d_{\text{str}} + \gamma, \quad \gamma \propto \mathcal{N}(0, \sigma_{\text{sens}}).$$

Due to the noise in the observations, the task may require several consecutive actions to finish. At every time step, the subject has three actions $a \in \{\text{present}, \text{absent}, \text{look}\}$ at their disposal: to make a decision that the signal is *present* or *absent* and to take another *look* at the signal. The role of the experimenter is to adjust the design $\mathbf{d} = \{d_{\text{str}}, d_{\text{obs}}\}$, where

d_{str} is the signal strength and d_{obs} is a discrete number of observations d_{obs} with the following design proposal distributions:

$$d_{\text{str}} \sim \text{Unif}(0, 4), \quad d_{\text{obs}} \sim \text{Unif}_{\text{discr}}(2, 10). \quad (22)$$

The subject can make such that the experiment will reveal characteristics of human behaviour. In particular, our goal is to identify the *hit* value parameter of the subject, which determines how much reward $r(a, s)$ the subject receives if the signal is both present and identified correctly:

$$r(a, s) = r_a(s) + r_{\text{step}},$$

$$r_a(s) = \theta_{\text{hit}}, \quad \text{when the signal is present and the action is } \textit{present}.$$

$$r_a(s) = 2, \quad \text{when the signal is absent and the action is } \textit{absent}.$$

$$r_a(s) = 0, \quad \text{when the action is } \textit{look}.$$

$$r_a(s) = -1, \quad \text{in other cases,}$$

where $r_{\text{step}} = -0.05$ is the constant cost of every consecutive action.

Models

In the context of a sequential signal detection task, we consider two models, each distinguished by their specific parameters:

$$\text{PPO} : x \sim PPO(\theta_{\text{hit}}, \theta_{\text{sens}}, \mathbf{d}), \quad (23)$$

$$\text{PR} : x \sim PR(\theta_{\text{hit}}, \theta_{\text{sens}}, \theta_{\text{low}}, \theta_{\text{len}}, \mathbf{d}). \quad (24)$$

The parameters of the models have the following priors:

$$\theta_{\text{hit}} \sim \text{Unif}(1, 7), \quad \theta_{\text{sens}} \sim \text{Unif}(0.1, 1), \quad (25)$$

$$\theta_{\text{low}} \sim \text{Unif}(0, 5), \quad \theta_{\text{len}} \sim \text{Unif}(0, 5). \quad (26)$$

In this study, both models are assumed to have a uniform prior distribution. The specific details of the individual models and their respective parameters are specified as follows.

PPO We implement the SDT task as an RL model due to the sequential nature of the task. In particular, the *look* action postpones the signal detection decision to the next observation. The model assumes that the subject acts according to the current observation o_{in} and an internal state β : $\pi(a | o_{\text{in}}, \beta)$. The internal state β is updated over trials by aggregating observations o_{in} using a *Kalman filter*, and after each trial, the agent chooses a new action.

As briefly discussed in Sect. 2, RL policies need to be retrained when their parameters change. To address this issue, the policy was parameterized and trained using a wide range of model parameters as policy inputs. This approach,

however, introduces a degree of model misspecification. While the PPO policy is inferred from training on a varied set of model parameters, synthetic participants utilise individual policies, each trained on distinct parameters. This discrepancy between the policy training and the behaviour of synthetic participants presents a realistic instance of model misspecification. Moreover, the PPO model, due to its inherent complexity and limited transparency, possesses a truly intractable likelihood. The resulting model was implemented using the PPO algorithm (Schulman et al., 2017).

PR An alternative to the PPO model is the PR model, which assumes sequential observations similar to the PPO model. In this model, a hypothesis test regarding the presence of a signal is performed after every observation, with the sequence of observations termed as *evidence* (Griffith et al., 2021).

A characteristic feature of the PR model is the calculation of the likelihood for the evidence, which is essentially a product of the likelihoods of each observation. While theoretically, the PR model could offer a likelihood, doing so is difficult in practice due to a sequential nature of the task. The model uses a *likelihood ratio*, denoted here as f_t , serving as a crucial decision variable. The evaluation of f_t against a specific threshold subsequently dictates the action a_t to be taken:

$$\begin{aligned} a_t &= \textit{present}, & \text{if } f_t &\leq \theta_{\text{low}}, \\ a_t &= \textit{absent}, & \text{if } f_t &\geq \theta_{\text{low}} + \theta_{\text{len}}, \\ a_t &= \textit{look}, & \text{if } \theta_{\text{low}} &\leq f_t \leq \theta_{\text{low}} + \theta_{\text{len}}. \end{aligned} \quad (27)$$

Here,

$$f_t = \prod_{i=1}^{d_{\text{obs}}} \frac{\omega_1}{\omega_2}, \quad \omega_1 \sim \mathcal{N}_{\text{CDF}}\left(\frac{1}{\theta_{\text{hit}} - 1}; d_{\text{str}}, \theta_{\text{sens}}\right), \quad (28)$$

$$\omega_2 \sim \mathcal{N}_{\text{CDF}}\left(\frac{1}{\theta_{\text{hit}} - 1}; 0, \theta_{\text{sens}}\right), \quad (29)$$

and $\mathcal{N}_{\text{CDF}}(\cdot; \mu, \nu)$ is the Gaussian CDF with the mean μ and standard deviation ν . For more information about the PR model, we refer the reader to Griffith et al. (2021).

Results

BOSMOS and MINEBED are the only methodologies capable of performing model selection in sequential signal detection models, as specified in Sect. 4.4, due to the intractability of their likelihoods. The experimental conditions are therefore very close to those in which these LFI approaches are usually applied, with the exception that we now know the ground-truth of synthetic participants for performance assessments.

BOSMOS showed a faster convergence of the estimates than MINEBED, requiring only 4 design trails to reduce the

majority of the uncertainty associated with model prediction accuracy and behaviour fitness error, as demonstrated in Fig. 3. In contrast, it took 20 design trials for MINEBED to converge, and extending it beyond 20 trials provided very little benefit. Similarly, as in the memory retention task from Sect. 4.3, the error in BOSMOS parameter estimates did not converge to zero, showing difficulty in predicting model parameters for PPO and PR models. Improving parameter inference may require modifying priors to encourage more diverse behaviours and selecting more descriptive experimental responses. Finally, BOSMOS outperformed MINEBED across all performance metrics after only one design trial, with the model predictive accuracy showing a

large difference, establishing BOSMOS as a clear favourite approach for this task.

An example of posterior distributions returned by BOSMOS is demonstrated in Fig. 4. Despite overall positive results, there are occasional cases in a population of synthetic participants where BOSMOS, along with MINEBED (as detailed in Appendix D), failed to converge on the ground-truth. These outliers may be attributed to the poor identifiability of the signal detection models, suggested earlier in the memory task, but also to the approximation inaccuracies accumulated over numerous trials. Further complications arise when models are misspecified, as this can increase the likelihood of particle collapse, thereby hindering particles from effectively

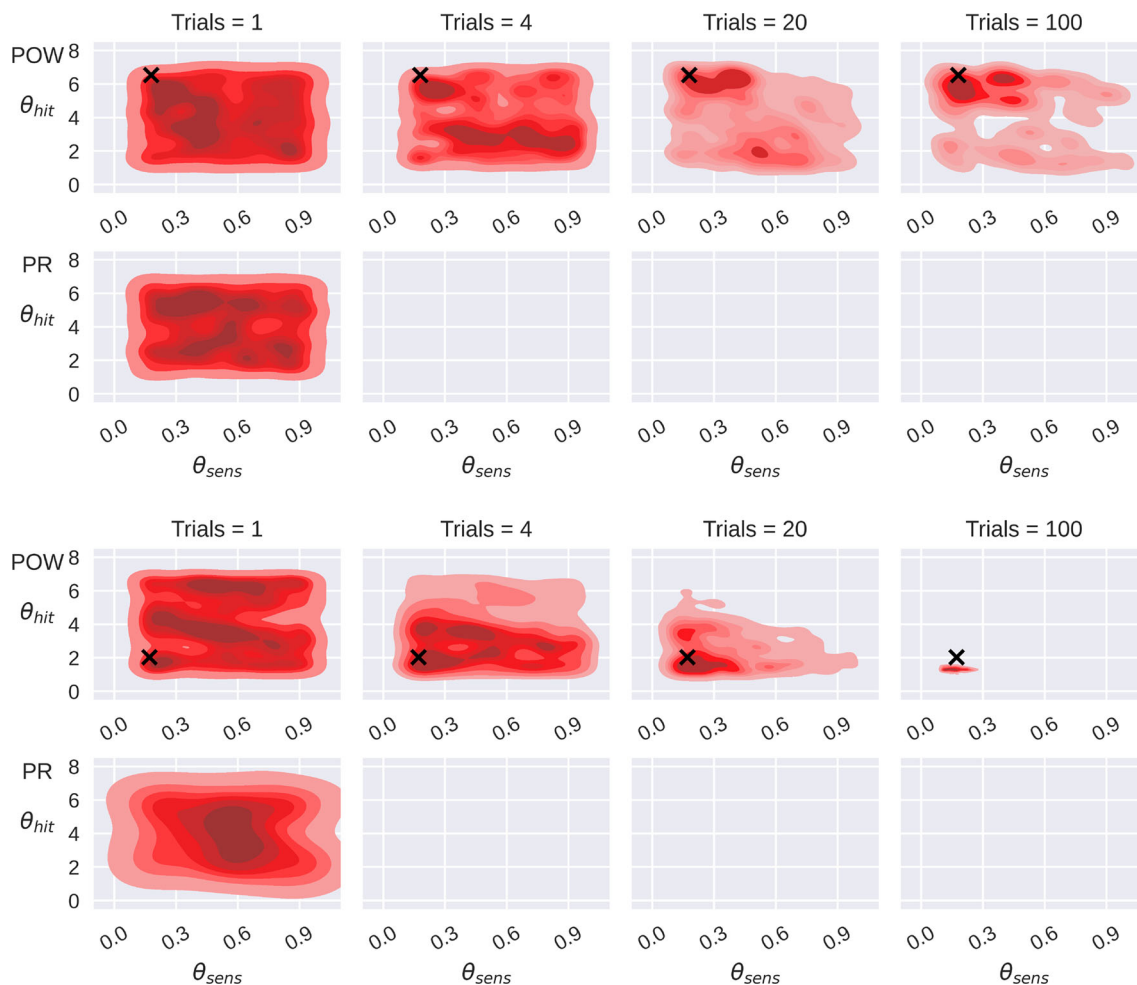


Fig. 4 An example of the evolution of the posterior approximation in each of the models tested resulting from BOSMOS in the signal detection task. The last bottom row panels are empty, as in both cases the posterior probability of the PR model becomes negligible, so that the particle approximation of this posterior does not contain any more particles. The true value of the parameters is indicated by the cross, and the true model is POW in both cases. BOSMOS successfully identified the ground-truth model in both cases: all posterior density (shaded area)

has been concentrated there by 20 trials, and no more particles exist in the other model. However, only in the first example (top panel) did the ground-truth parameter values (cross) fall inside the 90% confidence interval, indicating some inconsistency in terms of the posterior convergence towards the ground-truth. The axes correspond to the model parameters: sensor noise (x -axis) and hit value (y -axis); θ_{low} and θ_{len} of the PR model are omitted to simplify visualisation

exploring the posterior distribution. This issue further underscores the importance of adequately defining cognitive models to mitigate potential issues of poor identifiability, model misspecification, and particle collapse. However, given that both methods operate in an LFI setting, some inconsistency between replicating the target behaviour and converging to the ground-truth parameters is to be expected when the models are poorly identifiable.

Risky Choice

Risky choice problems are typical tasks used in psychology, cognitive science, and economics to study attitudes towards uncertainty. Specifically, risk refers to *quantifiable* uncertainty, where a decision-maker is aware of probabilities associated with different outcomes (Knight, 1985). In risky choice problems, individuals are presented with options that are lotteries (i.e., probability distributions of outcomes). For example, a risky choice problem could be a decision between winning 100 euros with a chance of 25%, or getting 25 euros with a chance of 99%. The choice is between two lotteries (100, 0.25; 0, 0.75) and (25, 0.99; 0, 0.01). The goal of the participant is to maximise the subjective reward of their single choice, so they need to assess the risk associated with outcomes in each lottery.

Several models have been proposed to explain tendencies in these tasks, including normative approaches derived from logic to descriptive approaches based on empirical findings (Johnson & Busemeyer, 2010). In this paper, we will consider four classic models (following Cavagnaro et al. (2013)): *expected utility* (EU) theory (Von Neumann & Morgenstern, 1990), *weighted expected utility* (WEU) theory (Hong, 1983), *original prospect theory* (OPT; (Kahneman & Tversky, 1979)) and *cumulative prospect theory* (CPT; (Tversky & Kahneman, 1992)). The risky choice models we consider consist of a subjective utility objective (characterising the amount of value an individual attaches to an outcome) and possibly a probability weighting function (reflecting the tendency for non-linear weighting of probabilities). Despite the long history of development, risky choices are still a focus of ongoing research (Begenau, 2020; Gächter et al., 2022; Frydman & Jin, 2022).

Task Description

Our objective is to maximise the reward obtained from risky choices. These choices typically comprise two or more alternatives, with each described by a set of outcome-probability pairs, in which probabilities sum up to one. While such problems in general may incorporate an endowment or entail

multiple stages, our model does not take these complexities into account in this iteration. The focus of our study is choice problems where individuals choose between two lotteries, denoted as A and B .

The design space for the risky-choice problems incorporates combined designs for both lotteries A and B : $\mathbf{d} = \{d_{plA}, d_{phA}, d_{plB}, d_{phB}\}$. Here, d_{phA} and d_{plA} represent probabilities of the high and low outcomes for the lottery A , and d_{phB} and d_{plB} analogously denote the same variables for the lottery B . We establish design proposal distributions for these variables as follows:

$$d_{plA} \sim \text{Unif}(0, 1), \quad d_{phA} \sim \text{Unif}(0, 1), \quad (30)$$

$$d_{plB} \sim \text{Unif}(0, 1), \quad d_{phB} \sim \text{Unif}(0, 1). \quad (31)$$

Please note that d_{pmA} and d_{pmB} can be analytically derived from $d_{pmA} = 2 - d_{plA} - d_{phA}$. Subsequently, the designs for each individual lottery ($d_{plA}, d_{pmA}, d_{phA}$) are normalised to sum up to one. Similar adjustments are made for the lottery B .

Taking into account the inherent variability of individual choices in risky problems, we assume that such choices are not deterministic (i.e., there is *choice stochasticity*). This assumption provides a likelihood for the ADO and LBIRD methods in our experiments. We adopt the definition provided by Cavagnaro et al. (2013) for the probability of choosing lottery A over B in a given choice problem i :

$$\phi_i(A_i | \theta_m, \epsilon) = \begin{cases} \epsilon, & \text{if } A_i \prec B_i \\ \frac{1}{2}, & \text{if } A_i \sim B_i \\ 1 - \epsilon, & \text{if } A_i \succ B_i \end{cases} \quad (32)$$

In this equation, θ_m refers to the model parameters, and ϵ is a value in the range $[0, 0.5]$ that quantifies the stochasticity of the choice. A zero ϵ signifies a deterministic choice. The preference for lottery A is assessed using the utility definitions distinct to each model.

Models

In our exploration of the risky choice task, we employ similar implementations as outlined by Cavagnaro et al. (2013) to examine four models:

$$\mathbf{EU} : x \sim EU(\theta_a, \theta_\epsilon, \mathbf{d}), \quad (35)$$

$$\mathbf{WEU} : x \sim WEU(\theta_x, \theta_y, \theta_\epsilon, \mathbf{d}), \quad (36)$$

$$\mathbf{OPT} : x \sim OPT(\theta_v, \theta_r, \theta_\epsilon, \mathbf{d}), \quad (37)$$

$$\mathbf{CPT} : x \sim CPT(\theta_v, \theta_r, \theta_\epsilon, \mathbf{d}). \quad (38)$$

Each of these models is characterised by a unique set of parameters, for which we use the following prior distributions:

$$\theta_a \sim \text{Unif}(0, 10), \quad \theta_v \sim \text{Unif}(0, 1), \quad (39)$$

$$\theta_r \sim \text{Unif}(0.01, 1), \quad \theta_x \sim \text{Unif}(-100, 0), \quad (40)$$

$$\theta_y \sim \text{Unif}(-100, 0), \quad \theta_\epsilon \sim \text{Unif}(0, 0.5), \quad (41)$$

For the purposes of consistency and simplicity, we assume a uniform prior distribution across all four models. The specifics of each model and their associated parameters are specified below.

EU Following Cavagnaro et al. (2013), we specify EU using indifference curves on the *Marschak-Machina* (MM) probability triangle. Lottery *A* consists of three outcomes (x_{lA} , x_{mA} , x_{hA}), and associated probabilities (p_{lA} , p_{mA} , p_{hA}). Lottery *A* can be represented using a right triangle (MM) with two of the probabilities as the plane (p_{lA} and p_{hA} as x and y axes, respectively). Hence, the design space for the lottery *A* consists of only the high and low probabilities (d_{plA} and d_{phA}). Lottery *B* can be represented on the triangle similarly (using d_{plB} and d_{phB}). Then, indifference curves can be drawn on this triangle, as their slope represents the marginal rate of substitution between the two probabilities. EU is defined using indifference curves that all have the same slope $\theta_a \in \theta_{EU}$. If the lottery *B* is riskier, $A \succ B$, if $|d_{phB} - d_{phA}| / |d_{plB} - d_{plA}| < \theta_a$. We refer the reader to Cavagnaro et al. (2013) for a more comprehensive explanation of this modelling approach.

WEU WEU is also defined using the MM-triangle, as per Cavagnaro et al. (2013). In contrast to EU, the slope of the indifference curves varies across the MM-triangle for WEU. This is achieved by assuming that all the indifference curves intersect at a point (θ_x, θ_y) outside the MM-triangle, where $[\theta_x, \theta_y] \in \theta_{WEU}$. Then, $A \succ B$, if $|d_{phA} - \theta_y| / |d_{plA} - \theta_x| > |d_{phB} - \theta_y| / |d_{plB} - \theta_x|$.

OPT In contrast to EU and WEU, OPT assumes that both the outcomes x and probabilities p have specific editing functions v and w , respectively. Assuming that for lottery *A*, $v(x_{low}^A) = 0$ and $v(x_{high}^A) = 1$, the utility objectives in OPT can be defined using $v(x_{middle}^A)$ as a parameter θ_v

$$u(A) = \begin{cases} w(d_{phA}) \cdot 1 + \theta_v \cdot (1 - w(d_{phA})), & \text{if } d_{plA} = 0 \\ w(d_{phA}) \cdot 1 + w(1 - d_{phA} - d_{plA}) \cdot \theta_v, & \text{otherwise.} \end{cases} \quad (42)$$

Utility $u(B)$ for lottery *B* can be calculated analogously, and $A_i \succ B_i$ if $u(A) > u(B)$. The probability weighting function $w(\cdot)$ used is the original work by Tversky and Kahneman (1992) is

$$w(p) = \frac{p^{\theta_r}}{(p^{\theta_r} + (1 - p)^{\theta_r})^{(1/\theta_r)}}, \quad (44)$$

where θ_r is a parameter describing the shape of the function. Thus, OPT has two parameters $[\theta_v, \theta_r] \in \theta_{OPT}$, describing the subjective utility of the middle outcome and the shape of the probability weighting function, respectively.

CPT CPT is defined similarly to OPT; however, the subjective utilities u for lottery *A* are calculated using

$$u(A) = w(d_{phA}) \cdot 1 + (w(1 - d_{plA}) - w(d_{phA})) \cdot \theta_v. \quad (45)$$

Utility $u(B)$ for lottery *B* is calculated similarly and $[\theta_v, \theta_r] \in \theta_{CPT}$.

Results

The risky choice task comprises four computational models, which significantly expand the space of models and make it much more computationally costly than previous tasks. Despite the larger model space, BOSMOS maintains its position as a preferred LFI approach to model selection, most notably when compared to the parameter estimation error of MINEBED from Fig. 2. With more models, BOSMOS's performance advantage over MINEBED grows, with BOSMOS exhibiting higher scalability for larger model spaces. Additionally, our experiments in Appendix C reveal that the choice of the Bayesian estimator, whether MAP or the *Bayesian information criterion*, has negligible impact on the overall performance of methods.

It is crucial to note that having several candidate models reduces model prediction accuracy by the LFI approaches; thus, we recommend reducing the number of candidate models as low as feasible. In terms of performance, BOSMOS is comparable to ground-truth likelihood approaches during the first four design trials, as shown in Fig. 3, since it is significantly easier to minimise uncertainty early in the trials. Similarly to the memory task, the error of LFI approximation becomes more apparent as the number of trials rises, as evidenced by comparing BOSMOS to ADO for the behavioural fitness error and model predictive accuracy. In terms of the parameter estimate error, BOSMOS performs marginally better than ADO.

Further investigation into the robustness of BOSMOS under model misspecification was evaluated through two additional risky choice experiments, as detailed in Appendix E. The first scenario, involving noise-induced misspecification in a risky choice task, showed that BOSMOS is capable of maintaining a significant level of performance even under moderate noise levels (0%-30%). The second scenario, involving an artificial parameter reduction within a PR model to simulate model misspecification, showed that BOSMOS could still achieve behavioural fitness convergence despite a decrease in available parameters. Please refer to Appendix E for more details.

Finally, BOSMOS has a relatively low runtime cost, especially compared to other methods (about 1 min per design trial). This brings adaptive model selection closer to being applicable to real-world experiments in risky choices. The proposed method can be useful in online experiments that include lag times between trials, for instance, in assessing investment decisions (e.g., Camerer (2004); Gneezy & Poters (1997)) or game-like settings (e.g., Bauckhage et al. (2012); Putkonen et al. (2022); Viljanen et al. (2017)) where the participant waits between events.

Discussion

In this paper, we proposed a simulator-based experimental design method for model selection, BOSMOS, that does design selection for model and parameter inference at a speed orders of magnitude higher than other methods, bringing the method closer to online design selection. This was made possible with the newly proposed approximation of the model likelihood and simulator-based utility objective. Despite needing orders of magnitude fewer simulations, BOSMOS significantly outperformed LFI alternatives in the majority of cases while being orders of magnitude faster, bringing the method closer to an online inference tool. Crucially, the time between experiment trials was reduced to less than a minute. Whereas in some settings this time between trials may be too long, BOSMOS is a viable tool in experiments where the tasks include a lag time, for instance, in studies of language learning (e.g., Gardner et al. (1997); Nioche et al. (2021)) and task interleaving (e.g., Payne et al. (2007); Brumby et al. (2009); Gebhardt et al. (2021); Katidioti et al. (2014)). Moreover, our code implementation represents a proof of concept and was not fully optimised for maximal efficiency: in particular, a parallel implementation that exploits multiple cores and batches of simulated experiments would enable additional speedups (Wu & Frazier, 2016). As an interactive and sample-efficient method, BOSMOS can help reduce the number of required experiments. This can be of interest to both the subject and the experimenter. In human trials, it allows for faster interventions (e.g., adjusting the treatment plan) in critical settings such as *intensive care units* or *randomised controlled trials*. However, it can also have detrimental applications, such as targeted advertising and collecting personal data; therefore, the principles and practises of responsible *artificial intelligence* (Dignum, 2019; Arrieta et al., 2020) also have to be taken into account in applying our methodology.

There are at least two remaining issues left for future work. The first issue we witnessed in our experiments is that the accuracy of behaviour imitation does not necessarily correlate with the convergence to ground-truth models. This usually happens due to poor identifiability in the model-

parameter space, which may be quite prevalent in current and future computational cognitive models since they are all designed to explain the same behaviour. Currently, the only way to address this problem is to use Bayesian approaches, such as BOSMOS, that quantify the uncertainty over the models and their parameters. The second issue is the consistency of the method: in selecting only the most informative designs, the methods may misrepresent the posterior and return an overconfident posterior. This bias may occur, for example, due to a poor choice of priors or summary statistics (Nunes & Balding, 2010; Fearnhead & Prangle, 2012) for the collected data (when the data is high-dimensional). Ultimately, these issues do not hinder the goal of automating experimental designs but introduce the necessity for a human expert, who would ensure that the uncertainty around estimated models is acceptable and the design space is sufficiently explored to make final decisions.

Future work for simulator-based model selection in computational cognitive science needs to consider adopting hierarchical models, accounting for the subjects' ability to adapt or change throughout the experiments, and incorporating amortised non-myopic design selection. A first step in this direction would be to study hierarchical models (Kim et al., 2014) which would allow adjusting prior knowledge for populations and expanding the theory development capabilities of model selection methods from a single individual to a group level. We could also remove the assumption on the stationarity of the model by proposing a dynamic model of subjects' responses that adapts to the history of previous responses and previous designs, which is more reasonable in longer settings of several dozens of trials. Lastly, amortised non-myopic design selections (Blau et al., 2022) would even further reduce the wait time between design proposals, as the model can be pre-trained before experiments, and would also improve design exploration by encouraging long-term planning of the experiments. Addressing these three potential directions may have a synergistic effect on each other, thus expanding the application of simulator-based model selection in cognitive science even further.

Supplementary Information

The article has the following accompanying supplementary materials:

- Appendix A shows the validity of the approximation of the entropy gain for the design selection rule;
- Appendix B details the algorithm for the proposed BOSMOS method and analyses its complexity;
- Appendix C contains tables with full experimental results, which shows additional design evaluation points;

- Appendix D showcases a side-by-side comparison of the posterior evolution resulting from BOSMOS and MINEBED for the signal detection task.
- Appendix E demonstrates additional experiments with BOSMOS in regards to model misspecification sensitivity.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s42113-023-00180-7>.

Funding Open Access funding provided by Aalto University. This work was supported by the Academy of Finland (Flagship programme: Finnish Center for Artificial Intelligence FCAI; grants 328400, 345604, 328400, 320181). AP and SC were funded by the Academy of Finland projects BAD (Project ID: 318559) and Human Automata (Project ID: 328813). AP was additionally funded by Aalto University School of Electrical Engineering. SC was also funded by Interactive Artificial Intelligence for Research and Development (AIRD) grant by Future Makers. SK was supported by the Engineering and Physical Sciences Research Council (EPSRC; Project ID: EP-W002973-1). Computational resources were provided by the Aalto Science-IT Project.

Availability of Data and Materials The paper uses simulated experiments which can be fully replicated with the code below.

Code Availability All code for replicating the experiments is available at <https://github.com/AaltoPML/BOSMOS>.

Declarations

Ethics Approval Not applicable

Consent to Participate Not applicable

Consent for Publication Not applicable

Conflict of Interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References


- Acerbi, L., Ma, W.J., & Vijayakumar, S. (2014). A framework for testing identifiability of Bayesian models of perception. *Advances in Neural Information Processing Systems*, 27
- Amin, H. U., & Malik, A. S. (2013). Human memory retention and recall processes. A review of EEG and fMRI studies. *Neurosciences*, 18(4), 330–44.
- Anderson, J. R. (1978). Arguments concerning representations for mental imagery. *Psychological Review*, 85(4), 249.
- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., et al. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115.
- Balandat, M., Karrer, B., Jiang, D., Daulton, S., Letham, B., Wilson, A. G., & Bakshy, E. (2020). BoTorch: A framework for efficient Monte-Carlo Bayesian optimization. *Advances in Neural Information Processing Systems*, 33, 21524–21538.
- Bauchhage, C., Kersting, K., Sifa, R., Thureau, C., Drachen, A., & Canossa, A. (2012). How players lose interest in playing a game: An empirical study based on distributions of total playing times. *2012 IEEE conference on computational intelligence and games (CIG)* (pp. 139–146)
- Beaumont, M. A., Zhang, W., & Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, 162(4), 2025–2035.
- Begenau, J. (2020). Capital requirements, risk choice, and liquidity provision in a business-cycle model. *Journal of Financial Economics*, 136(2), 355–378.
- Blau, T., Bonilla, E.V., Chades, I., & Dezfouli, A. (2022). Optimizing sequential experimental design with deep reinforcement learning. *International conference on machine learning* (pp. 2107–2128)
- Boelts, J., Lueckmann, J.-M., Gao, R., & Macke, J. H. (2022). Flexible and efficient simulation-based inference for models of decision-making. *eLife*, 11, e77220.
- Brumby, D.P., Salvucci, D.D., & Howes, A. (2009). Focus on driving: How cognitive constraints shape the adaptation of strategy when dialing while driving. *Proceedings of the SIGCHI conference on human factors in computing systems* (p. 1629–1638). New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/1518701.1518950>
- Camerer, C.F. (2004). Prospect theory in the wild: Evidence from the field. *Advances in Behavioral Economics*, 148–161
- Cavagnaro, D. R., Gonzalez, R., Myung, J. I., & Pitt, M. A. (2013). Optimal decision stimuli for risky choice experiments: An adaptive approach. *Management Science*, 59(2), 358–375.
- Cavagnaro, D. R., Myung, J. I., Pitt, M. A., & Kujala, J. V. (2010). Adaptive design optimization: A mutual information-based approach to model discrimination in cognitive science. *Neural Computation*, 22(4), 887–905.
- Cavagnaro, D. R., Pitt, M. A., Gonzalez, R., & Myung, J. I. (2013). Discriminating among probability weighting functions using adaptive design optimization. *Journal of Risk and Uncertainty*, 47(3), 255–289.
- Cranmer, K., Brehmer, J., & Louppe, G. (2020). The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences*, 117(48), 30055–30062.
- Del Moral, P., Doucet, A., & Jasra, A. (2006). Sequential Monte-Carlo samplers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3), 411–436.
- Dignum, V. (2019). *Responsible artificial intelligence: How to develop and use AI in a responsible way*. Springer Nature
- Doucet, A., Godsill, S., & Andrieu, C. (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and computing*, 10, 197–208.
- Fearnhead, P., & Prangle, D. (2012). Constructing summary statistics for approximate Bayesian computation: Semi-automatic approximate Bayesian computation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(3), 419–474.
- Foster, A., Ivanova, D.R., Malik, I., & Rainforth, T. (2021). Deep adaptive design: Amortizing sequential Bayesian experimental design. *International conference on machine learning* (pp. 3384–3395)

- Frazier, P.I. (2018). Bayesian optimization. Recent advances in optimization and modeling of contemporary problems (pp. 255–278). *Informatics*
- Frydman, C., & Jin, L. J. (2022). Efficient coding and risky choice. *The Quarterly Journal of Economics*, 137(1), 161–213.
- Gächter, S., Johnson, E. J., & Herrmann, A. (2022). Individual-level loss aversion in riskless and risky choices. *Theory and Decision*, 92(3), 599–624.
- Gardner, R. C., Tremblay, P. F., & Masgoret, A.-M. (1997). Towards a full model of second language learning: An empirical investigation. *The Modern Language Journal*, 81(3), 344–362.
- Gebhardt, C., Oulasvirta, A., & Hilliges, O. (2021). Hierarchical reinforcement learning explains task interleaving behavior. *Computational Brain & Behavior*, 4(3), 284–304.
- Gneezy, U., & Potters, J. (1997). An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics*, 112(2), 631–645.
- Greenhill, S., Rana, S., Gupta, S., Vellanki, P., & Venkatesh, S. (2020). Bayesian optimization for adaptive experimental design: A review. *IEEE Access*, 8, 13937–13948.
- Griffith, T., Baker, S.-A., & Lepora, N. F. (2021). The statistics of optimal decision-making: Exploring the relationship between signal detection theory and sequential analysis. *Journal of Mathematical Psychology*, 103, 102544.
- Gutmann, M.U., & Corander, J. (2016). Bayesian optimization for likelihood-free inference of simulator-based statistical models. *Journal of Machine Learning Research*
- Hong, C.S. (1983). A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox. *Econometrica*, 51(4), 1065–1092. Retrieved 2022-09-27, from <http://www.jstor.org/stable/1912052>
- Howes, A., Lewis, R. L., & Vera, A. (2009). Rational adaptation under task and processing constraints: Implications for testing theories of cognition and action. *Psychological Review*, 116(4), 717.
- Ivanova, D. R., Foster, A., Kleinegesse, S., Gutmann, M. U., & Rainforth, T. (2021). Implicit deep adaptive design: Policy-based experimental design without likelihoods. *Advances in Neural Information Processing Systems*, 34, 25785–25798.
- Johnson, J. G., & Busemeyer, J. R. (2010). Decision-making under risk and uncertainty. *WIREs Cognitive Science*, 1(5), 736–749. <https://doi.org/10.1002/wcs.76>
- Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4, 237–285.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–292.
- Kangasrääsio, A., Jokinen, J. P., Oulasvirta, A., Howes, A., & Kaski, S. (2019). Parameter inference for computational cognitive models with approximate Bayesian computation. *Cognitive Science*, 43(6), e12738.
- Katidioti, I., Borst, J. P., & Taatgen, N. A. (2014). What happens when we switch tasks: Pupil dilation in multitasking. *Journal of Experimental Psychology: Applied*, 20(4), 380.
- Kiefer, J. (1959). Optimum experimental designs. *Journal of the Royal Statistical Society: Series B (Methodological)*, 21(2), 272–304.
- Kim, W., Pitt, M. A., Lu, Z.-L., Steyvers, M., & Myung, J. I. (2014). A hierarchical adaptive approach to optimal experimental design. *Neural Computation*, 26(11), 2465–2492.
- Kingma, D.P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980)
- Kleinegesse, S. & Gutmann, M.U. (2020). Bayesian experimental design for implicit models by mutual information neural estimation. *International conference on machine learning* (pp. 5316–5326)
- Kleinegesse, S., & Gutmann, M.U. (2021). Gradient-based Bayesian experimental design for implicit models using mutual information lower bounds. arXiv preprint [arXiv:2105.04379](https://arxiv.org/abs/2105.04379)
- Knight, F. H. (1985). *Risk, uncertainty and profit* (Repr). Chicago: University of Chicago Press.
- Kong, A., Liu, J. S., & Wong, W. H. (1994). Sequential imputations and Bayesian missing data problems. *Journal of the American statistical association*, 89(425), 278–288.
- Lee, M. D., Criss, A. H., Devezer, B., Donkin, C., Etz, A., Leite, F. P., et al. (2019). Robust modeling in cognitive science. *Computational Brain & Behavior*, 2(3), 141–153.
- Lindley, D. V. (1956). On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, 27(4), 986–1005.
- Liu, J. S., & Chen, R. (1998). Sequential Monte Carlo methods for dynamic systems. *Journal of the American Statistical Association*, 93(443), 1032–1044.
- Madsen, J. K., Bailey, R., Carrella, E., & Koralus, P. (2019). Analytic versus computational cognitive models: Agent-based modeling as a tool in cognitive sciences. *Current Directions in Psychological Science*, 28(3), 299–305.
- Madsen, J. K., Bailey, R. M., & Pilditch, T. D. (2018). Large networks of rational agents form persistent echo chambers. *Scientific Reports*, 8(1), 1–8.
- Marin, J.-M., Pudlo, P., Robert, C. P., & Ryder, R. J. (2012). Approximate Bayesian computational methods. *Statistics and computing*, 22(6), 1167–1180.
- Michaud, I.J. (2019). Simulation-based Bayesian experimental design using mutual information. Dissertation, North Carolina State University.
- Moon, H.-S., Do, S., Kim, W., Seo, J., Chang, M., & Lee, B. (2022). Speeding up inference with user simulators through policy modulation. *CHI conference on human factors in computing systems* (pp. 1–21)
- Musso, C., Oudjane, N., & Le Gland, F. (2001). Improving regularised particle filters. In Doucet, A., de Freitas, N., Gordon, N. (eds.). *Sequential Monte Carlo methods in practice, Statistics for Engineering and Information Science*. Springer, New York, NY (pp. 247–271)
- Myung, J. I., Cavagnaro, D. R., & Pitt, M. A. (2013). A tutorial on adaptive design optimization. *Journal of Mathematical Psychology*, 57(3–4), 53–67.
- Nioche, A., Murena, P.-A., de la Torre-Ortiz, C., & Oulasvirta, A. (2021). Improving artificial teachers by considering how people learn and forget. *26th international conference on intelligent user interfaces* (pp. 445–453)
- Nunes, M.A., & Balding, D.J. (2010). On optimal selection of summary statistics for approximate Bayesian computation. *Statistical Applications in Genetics and Molecular Biology*, 9(1)
- Oulasvirta, A., Jokinen, J.P., & Howes, A. (2022). Computational rationality as a theory of interaction. *CHI conference on human factors in computing systems* (pp. 1–14). ACM
- Overstall, A. M., & Woods, D. C. (2017). Bayesian design of experiments using approximate coordinate exchange. *Technometrics*, 59(4), 458–470.
- Papamakarios, G., Sterratt, D., & Murray, I. (2019). Sequential neural likelihood: Fast likelihood-free inference with autoregressive flows. *The 22nd international conference on artificial intelligence and statistics* (pp. 837–848)
- Payne, S. J., Duggan, G. B., & Neth, H. (2007). Discretionary task interleaving: Heuristics for time allocation in cognitive foraging. *Journal of Experimental Psychology: General*, 136(3), 370.
- Peterson, W., Birdsall, T., & Fox, W. (1954). The theory of signal detectability. *Transactions of the IRE Professional Group on Information Theory*, 4(4), 171–212. <https://doi.org/10.1109/TIT.1954.1057460>

- Pudlo, P., Marin, J.-M., Estoup, A., Cornuet, J.-M., Gautier, M., & Robert, C. P. (2016). Reliable ABC model choice via random forests. *Bioinformatics*, 32(6), 859–866.
- Putkonen, A., Nioche, A., Tanskanen, V., Klami, A., & Oulasvirta, A. (2022). How suitable is your naturalistic dataset for theory-based user modeling? *Proceedings of the 30th ACM conference on user modeling, adaptation and personalization* (pp. 179–190)
- Rasmussen, C.E. (2004). Gaussian processes in machine learning. In: Bousquet, O., von Luxburg, U., & Rätsch, G. (eds). *Advanced Lectures on Machine Learning*. ML 2003. Lecture Notes in Computer Science, vol 3176. Springer, Berlin, Heidelberg. (pp. 63–71)
- Robert, C.P. (2007). *The Bayesian choice: From decision-theoretic foundations to computational implementation* (Vol. 2). Springer
- Rubin, D.C., Hinton, S., & Wenzel, A. (1999). The precise time course of retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(5), 1161–1176. Retrieved 2022-10-26, from <https://doi.org/10.1037/0278-7393.25.5.1161>
- Rubin, D.C., & Wenzel, A.E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, 103(4), 734–760. Retrieved 2022-10-26, from <https://doi.org/10.1037/0033-295X.103.4.734>
- Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. arXiv. Retrieved from [arXiv:1707.06347](https://arxiv.org/abs/1707.06347). <https://doi.org/10.48550/ARXIV.1707.06347>
- Seeger, M. (2004). Gaussian processes for machine learning. *International Journal of Neural Systems*, 14(02), 69–106.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379–423.
- Sisson, S. A., Fan, Y., & Beaumont, M. (2018). *Handbook of approximate Bayesian computation*. CRC Press.
- Srinivas, N., Krause, A., Kakade, S.M., & Seeger, M. (2009). Gaussian process optimization in the bandit setting: No regret and experimental design. arXiv preprint [arXiv:0912.3995](https://arxiv.org/abs/0912.3995)
- Sutton, R.S., & Barto, A.G. (2018). *Reinforcement learning: An introduction*. MIT Press.
- Swets, J.A., Tanner, Jr. W.P., & Birdsall, T.G. (1961). Decision processes in perception. *Psychological Review*, 68(5), 301–340. Retrieved 2022-10-26, <https://doi.org/10.1037/h0040547>
- Tanner, W.P., & Swets, J.A. (1954). A decision-making theory of visual detection. *Psychological Review*, 61(6), 401–409. Retrieved 2022-10-26, from <https://doi.org/10.1037/h0058700>
- Tauber, S., Navarro, D. J., Perfors, A., & Steyvers, M. (2017). Bayesian models of cognition revisited: Setting optimality aside and letting data drive psychological theory. *Psychological Review*, 124(4), 410–441.
- Tavaré, S., Balding, D. J., Griffiths, R. C., & Donnelly, P. (1997). Inferring coalescence times from DNA sequence data. *Genetics*, 145(2), 505–518.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323. https://doi.org/10.1007/978-3-319-20451-2_24
- Valentin, S., Kleinegesse, S., Bramley, N.R., Gutmann, M., & Lucas, C. (2021). Bayesian experimental design for intractable models of cognition. *Proceedings of the annual meeting of the Cognitive Science Society* (Vol. 43)
- Viljanen, M., Airola, A., Heikkonen, J., & Pahikkala, T. (2017). Play-time measurement with survival analysis. *IEEE Transactions on Games*, 10(2), 128–138.
- Von Neumann, J., & Morgenstern, O. (1990). 3. The notion of utility. *Theory of games and economic behavior* (3rd ed., pp. 15–31). Princeton, New Jersey: Princeton University Press. (original date: 1944)
- Wickens, T.D. (2002). *Elementary signal detection theory*. Oxford ; New York: Oxford University Press
- Wixted, J.T. (2020). The forgotten history of signal detection theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(2), 201–233. Retrieved 2022-10-26, from <https://doi.org/10.1037/xlm0000732>
- Wu, J., & Frazier, P. (2016). The parallel knowledge gradient method for batch Bayesian optimization. *Advances in Neural Information Processing Systems*, 29
- Yang, J., & Qiu, W. (2005). A measure of risk and a decision-making model based on expected utility and entropy. *European Journal of Operational Research*, 164(3), 792–799.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Alexander Aushev¹  · Aini Putkonen² · Grégoire Clarté³ · Suyog Chandramouli² · Luigi Acerbi³ · Samuel Kaski^{1,4} · Andrew Howes⁵

Aini Putkonen
aini.putkonen@aalto.fi

Grégoire Clarté
gregoire.clarte@helsinki.fi

Suyog Chandramouli
suyog.chandramouli@aalto.fi

Luigi Acerbi
luigi.acerbi@helsinki.fi

Samuel Kaski
samuel.kaski@aalto.fi

¹ Department of Computer Science, Aalto University, Espoo 02150, Finland

² Department of Information and Communications Engineering, Aalto University, Espoo 02150, Finland

³ Department of Computer Science, University of Helsinki, Helsinki 00560, Finland

⁴ Department of Computer Science, University of Manchester, Manchester M13 9PR, UK

⁵ School of Computer Science, University of Birmingham, Birmingham B15 2TT, UK