Maximin Robust Bayesian Experimental Design

Hany Abdulsamad
UvA–Bosch Delta Lab
University of Amsterdam
h.abdulsamad@uva.nl

Sahel Iqbal
Department of Statistics
University of Oxford
sahel.iqbal@stats.ox.ac.uk

Christian A. Naesseth
UvA–Bosch Delta Lab
University of Amsterdam
c.a.naesseth@uva.nl

Takuo Matsubara
Business School
University of Sydney
takuo.matsubara@sydney.edu.au

Adrien Corenflos
Department of Statistics
University of Warwick
adrien.corenflos@warwick.ac.uk

Abstract

We address the brittleness of Bayesian experimental design under model misspecification by formulating the problem as a max–min game between the experimenter and an adversarial nature subject to information-theoretic constraints. We demonstrate that this approach yields a robust objective governed by Sibson's α-mutual information (MI), which identifies the α-tilted posterior as the robust belief update and establishes the Rényi divergence as the appropriate measure of conditional information gain. To mitigate the bias and variance of nested Monte Carlo estimators needed to estimate Sibson's α-MI, we adopt a PAC-Bayes framework to search over stochastic design policies, yielding rigorous high-probability lower bounds on the robust expected information gain that explicitly control finite-sample error.

1 Introduction

Experimental design is a foundational component of scientific and engineering inquiry, enabling efficient knowledge acquisition across a wide range of domains including physics, epidemiology, neuroscience, biology, and robotics [Melendez et al., 2021, Cook et al., 2008, Shababo et al., 2013, Liepe et al., 2013, Schultheis et al., 2020]. In such settings, each experiment can involve substantial costs related to time, labor, materials, and financial investments, making it imperative to extract as much information as possible from every trial. By explicitly modeling uncertainty over latent parameters and quantifying the informativeness of potential measurements, Bayesian experimental design [Lindley, 1956, Chaloner and Verdinelli, 1995] offers a principled framework for selecting experiments that are maximally informative. To do so, it ties decision-making and Bayesian inference into a single objective aiming to find a design minimizing posterior distribution *uncertainty*, measured in terms of entropy and averaged over the marginal distribution of prospective measurements. This information-theoretic approach enables efficient allocation of experimental effort, reducing both cost and risk while accelerating scientific discovery, particularly where simulators enable large-scale evaluation of designs prior to deployment.

However, accurate simulators remain the exception rather than the rule, even as modeling capabilities continue to improve. Mechanical, chemical, and biological systems often exhibit complex and only partially understood interactions that cannot be faithfully captured by idealized assumptions [Rondelez, 2012, van Sluijs et al., 2024]. Designs optimized under misspecified simulators may therefore lead to ineffective experiments or, in extreme cases, counterproductive outcomes. Hence, when the model is misspecified, robustness is paramount both in the posterior update and in the data generation governed by the marginal distribution. The decision-making literature has a rich history of addressing misspecification in the data-generating process [Whittle, 1990, Petersen et al., 2000], while the

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

Bayesian inference literature has developed extensive methodology for better posterior updates under likelihood misspecification [Grünwald, 2012, Bissiri et al., 2016, Ghosh and Basu, 2016, Holmes and Walker, 2017, Grünwald and van Ommen, 2017]. The central challenge in robust Bayesian experimental design is to unify these perspectives in a single framework that treats robustness in data generation and inference jointly, capturing their interaction rather than addressing them in isolation.

This problem has attracted increasing attention recently. One popular approach to robustify Bayesian inference is to replace the log-likelihood with a generalized utility yielding what are known as Gibbs posteriors [Zhang, 2006, Bissiri et al., 2016]. Recent contributions to robust Bayesian experimental design have adopted this framework, substituting standard posteriors appearing in the expected information gain objective with Gibbs posteriors. Additionally, these methods modify the generative component, either by introducing user-specified distributions [Overstall et al., 2025] or by constructing simulators implicitly defined by the chosen scoring rule [Barlas et al., 2025]. While these approaches offer practical mechanisms for robustifying experimental design, their *ad hoc* formulations lack a coherent optimization perspective of misspecification.

In contrast, we propose a principled framework for robust Bayesian experimental design under model misspecification, formulated as a max–min statistical decision problem [Berger, 1985]. Misspecification is handled through the lens of distributionally robust optimization [Kuhn et al., 2025] by considering worst-case perturbations of the data-generating process within a Kullback–Leibler neighborhood centered at a nominal model representing the experimenter's best approximation to reality. This formulation yields a tractable solution in the form of a Rényi mutual information, specifically Sibson's α-mutual information [Sibson, 1969, Verdú, 2015, Esposito et al., 2022], which we interpret as a robust expected information gain. However, this quantity must typically be estimated using nested Monte Carlo estimators, which yield biased and stochastic evaluations and thereby undermine the applicability of standard deterministic optimization methods. We therefore adopt a PAC-Bayes approach [McAllester, 1998] that optimizes randomized design policies and provide high-probability guarantees on the true robust expected information gain.

This article is organized as follows: Section 2 reviews Bayesian experimental design; Section 3 formulates robust experimental design as a max–min problem, yielding a robust expected information gain based on Sibson's α-mutual information; Section 4 develops a nested Monte Carlo estimator and analyzes its bias and concentration; Section 5 establishes a PAC-Bayes framework for robust design optimization with finite-sample guarantees; Section 6 discusses related work; Section 7 reports numerical evaluations. Proofs, technical assumptions, and other details are deferred to the appendix.

2 Background

Let Θ, Ξ and $\mathcal{X}$ denote the sets of parameters, designs, and outcomes of an experiment, respectively. We write $\mathcal{P}(W)$ for the set of all probability densities defined on a set W. In Bayesian experimental design, an experimenter selects a design $\xi \in \Xi$ with the objective of inferring latent parameters $\theta \in \Theta$ from observed measurements $x \in \mathcal{X}$. The experimenter begins with a prior belief $p(\theta)$ over the parameters and assumes a conditional likelihood function $p(x \mid \theta, \xi)$, which specifies how the design ξ relates the latent parameter θ to the observable outcome x. The induced marginal distribution of the measurement is $p(x \mid \xi)$ and the Bayes posterior is $p(\theta \mid x, \xi) \propto p(\theta)\, p(x \mid \theta, \xi)$.

Expected Information Gain. Bayesian experimental design admits a decision-theoretic formulation through the framework of *scoring rules* [Gneiting and Raftery, 2007]. For a latent parameter θ, the experimenter issues a probabilistic forecast $r \in \mathcal{P}(\Theta)$. The quality of this forecast is evaluated by a scoring rule $S\colon \mathcal{P}(\Theta) \times \Theta \to \mathbb{R}$, where $S(r, \theta)$ denotes the utility of reporting the density r when the true parameter is θ [Bernardo and Smith, 1994]. A *proper* scoring rule is one that incentivizes the experimenter to report their true belief: if the experimenter's belief is $\omega \in \mathcal{P}(\Theta)$, then the expected utility is maximized by reporting exactly $r = \omega$ [Savage, 1971, Gneiting and Raftery, 2007], so that:

$$\omega \in \operatorname*{arg\,max}_{r \in \mathcal{P}(\Theta)} \int S(r, \theta)\, \omega(\theta)\, \mathrm{d}\theta. \tag{1}$$

If the maximizer is unique, the scoring rule is *strictly* proper. Before observing data, the experimenter's forecast is the prior density $p(\theta)$. After observing x under design ξ, the experimenter may issue a refined forecast. We represent a data-dependent forecast rule by the map $\psi_\xi\colon \mathcal{X} \to \mathcal{P}(\Theta)$. The value

of an experiment is the optimal post-observation score improvement over the prior forecast:

$$I(\xi) \coloneqq \int p(x \mid \xi) \left[\sup_{r_x \in \mathcal{P}(\Theta)} \int S(r_x, \theta)\, p(\theta \mid x, \xi)\, \mathrm{d}\theta \right] \mathrm{d}x - \int S(p, \theta)\, p(\theta)\, \mathrm{d}\theta$$
$$= \sup_{\psi_\xi : \mathcal{X} \to \mathcal{P}(\Theta)} \int p(\theta, x \mid \xi) \left[S(\psi_\xi(x), \theta) - S(p, \theta) \right] \mathrm{d}\theta\, \mathrm{d}x.$$

Only the post-observation forecast rule $\psi_\xi(x)$ is subject to optimization here, while the pre-observation forecast $p(\theta)$ is a fixed reference against which the informativeness of ξ is measured. By (1), given a proper scoring rule, the supremum of this objective is attained at $\psi^\star_\xi(x) = p(\cdot \mid x, \xi)$.

This construction may appear unnecessary at first glance. The fact that the experimenter's optimal forecast is the standard Bayes posterior is intuitive. This formulation, however, becomes advantageous when the joint law $p(\theta, x \mid \xi)$ is misspecified, as we discuss in Section 3.

Under the strictly proper logarithmic scoring rule $S_{\log}(r, \theta) \coloneqq \log r(\theta)$, the value $I(\xi)$ recovers Shannon's expected information gain [EIG; Lindley, 1956, Bernardo, 1979], that is $\int p(\theta, x \mid \xi) \big[\log p(\theta \mid x, \xi) - \log p(\theta) \big] \mathrm{d}\theta\, \mathrm{d}x$, the mutual information between θ and x under ξ. This objective has seen wide adoption in modern Bayesian experimental design [BED; Chaloner and Verdinelli, 1995, Rainforth et al., 2024].

Definition 1 (Expected information gain)**.** *The expected information gain of a design ξ, defined as the mutual information between the parameters θ and outcomes x, admits several representations:*

$$I(\xi) \coloneqq I(\theta; x)(\xi) = \mathbb{E}_{p(x \mid \xi)} \Big[\mathbb{H}\left[p(\theta)\right] - \mathbb{H}\left[p(\theta \mid x, \xi)\right] \Big]$$
$$= \mathbb{E}_{p(x \mid \xi)} \Big[\mathbb{D}_{\mathrm{KL}} \big[p(\theta \mid x, \xi) \;\big\|\; p(\theta) \big] \Big] \tag{2a}$$
$$= \mathbb{E}_{p(\theta)} \Big[\mathbb{D}_{\mathrm{KL}} \big[p(x \mid \theta, \xi) \;\big\|\; p(x \mid \xi) \big] \Big] = I(x; \theta)(\xi) \tag{2b}$$

where $\mathbb{H}$ denotes Shannon's differential entropy and $\mathbb{D}_{\mathrm{KL}}$ the Kullback–Leibler divergence.

Although Shannon's mutual information is symmetric, so that $I(\theta; x) \equiv I(x; \theta)$, the two orderings admit distinct interpretations. The ordering $I(\theta; x)$, given by (2a), reflects the inferential perspective, measuring the expected relative entropy from posterior to the prior of the latent parameter θ, after observing data x. In contrast, the ordering $I(x; \theta)$, as expressed in (2b), offers a predictive interpretation, quantifying the extent to which knowledge of θ explains the variability of the outcomes x. While these viewpoints coincide numerically for Shannon's mutual information, the distinction is important when considering general information measures for which symmetry need not hold.

Given this definition, the goal of Bayesian experimental design is therefore to select the designs $\xi^\star$ that maximize the expected information gain: $\xi^\star = \arg\max_{\xi \in \Xi} I(\xi)$.

Stochastic Design Policies. Standard experimental design treats the expected information gain as a function of deterministic designs $\xi \in \Xi$. This approach is adequate provided an exact and deterministic oracle of the objective is available. However, the expected information gain is often intractable and must be approximated by nested Monte Carlo estimators [Rainforth et al., 2018], yielding *biased* and *noisy* oracles. To address this challenge, we turn to a PAC-Bayes framework, which provides high-probability guarantees on the estimation error but requires modeling designs as random variables [Flynn et al., 2023, Alquier, 2024].

Rather than committing to a single design ξ, we allow the experimenter to specify a probability distribution over the design space, denoted by $\pi \in \Pi$, referred to as a stochastic *policy*. Concretely, given a generative process $p(\theta, x, \xi) = p(\theta)\, p(x \mid \theta, \xi)\, \pi(\xi)$, the expected information gain associated with a policy π is defined as the *conditional* mutual information between x and θ given ξ:

$$I(\pi) \coloneqq \mathbb{D}_{\mathrm{KL}} \big[p(\theta, x, \xi) \;\big\|\; p(\theta)\, p(x \mid \xi)\, \pi(\xi) \big] = \mathbb{E}_\pi \Big[\mathbb{D}_{\mathrm{KL}} \big[p(\theta, x \mid \xi) \;\big\|\; p(\theta)\, p(x \mid \xi) \big] \Big].$$

The design optimization problem is reframed as $\pi^\star = \arg\max_{\pi \in \Pi} I(\pi)$. Beyond satisfying PAC-Bayes requirements, this relaxation allows us to leverage a rich literature of stochastic policy optimization techniques [Peters et al., 2010, Haarnoja et al., 2018], which have shown strong performance in amortizing sequential Bayesian experimental design [Blau et al., 2022, Iqbal et al., 2024].

3 Robust Experimental Design

When the measurements encountered at deployment are indeed generated from the assumed conditional likelihood $p(x \mid \theta^\star, \xi)$ for some $\theta^\star \in \Theta$, Bernardo [1979] provides a decision-theoretic justification for the expected information gain. In this well-specified regime, the EIG faithfully reflects the anticipated value of an experiment.

However, in many applications the likelihood is only a modeling surrogate and may differ substantially from the true data-generating process. Such discrepancies are a source of systematic error in statistical inference under model misspecification [Grünwald and van Ommen, 2017]. The expected information gain is particularly fragile because of its nested dependence on the likelihood: it determines both the posterior $p(\theta \mid x, \xi) \propto p(\theta)\, p(x \mid \theta, \xi)$ and the marginal $p(x \mid \xi) = \int p(x \mid \theta, \xi)\, p(\theta)\, \mathrm{d}\theta$ that defines the outer expectation. Given misspecification, $I(\xi)$ may become an unreliable utility, favoring designs that appear highly informative under the assumed model but fail to deliver the anticipated information gain when evaluated against the true data-generating process.

Maximin Experimental Design. To treat this issue systematically, we cast experimental design under misspecification as a two-player zero-sum Stackelberg game between the experimenter and *nature* [von Neumann and Morgenstern, 1944, Başar and Olsder, 1998, Grünwald and Dawid, 2004]. Nature selects a joint law $q \colon \Xi \to \mathcal{P}(\Theta \times \mathcal{X})$ from an admissible set $\mathcal{Q}$ as the data-generating process and the experimenter counters with a design policy $\pi \in \Pi$. By perturbing the full joint $q(\theta, x \mid \xi)$, nature is allowed to represent scenarios of both prior miscalibration and likelihood misspecification.

We assume the experimenter knows the admissible set $\mathcal{Q}$ but not nature's specific choice $q \in \mathcal{Q}$, whereas nature observes the policy $\pi \in \Pi$ but not the experimenter's realized design draw ξ. The experimenter evaluates a policy via its expected worst-case payoff $V(\pi, q) \coloneqq \mathbb{E}_\pi\big[U(\xi, q)\big]$, and seeks robustness against nature's choice by solving the maximin problem [Wald, 1950, Berger, 1985]

$$\pi^\star \coloneqq \arg\sup_{\pi \in \Pi} \inf_{q \in \mathcal{Q}} V(\pi, q).$$

To define the pointwise payoff $U(\xi, q)$, we return to the decision-theoretic formulation of Section 2. Assuming the experimenter commits to the prior $p(\theta)$ as a fixed reference forecast, the value of a design ξ against an adversarial joint law $q(\theta, x \mid \xi)$ can be expressed as:

$$U(\xi, q) \coloneqq \sup_{\psi_\xi : \mathcal{X} \to \mathcal{P}(\Theta)} \int q(\theta, x \mid \xi) \Big[\log \psi_\xi(x)(\theta) - \log p(\theta) \Big] \, \mathrm{d}\theta \, \mathrm{d}x.$$

For any fixed $q(\theta, x \mid \xi)$, the supremum is attained at $\psi^\star_\xi(x) = q(\cdot \mid x, \xi)$ and the payoff reduces to:

$$U(\xi, q) = \mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \,\big\|\, p(\theta)\, q(x \mid \xi) \big].$$

This quantity is not, in general, the standard mutual information under q, but rather a divergence from $q(\theta, x \mid \xi)$ to a mixed law combining the experimenter's prior $p(\theta)$ with nature's marginal $q(x \mid \xi)$.

When the set $\mathcal{Q}$ is unrestricted, the worst-case formulation is vacuous, since nature can trivially drive the information gain to zero by enforcing independence between θ and x. A meaningful notion of robustness therefore requires $\mathcal{Q}$ to constrain nature's flexibility by specifying how far $q(\theta, x \mid \xi)$ may deviate from the nominal reference $p(\theta, x \mid \xi)$, which encodes the experimenter's best guess of the true data-generating process. Because the experimenter commits to a stochastic policy, it is natural to specify this restriction with an average misspecification budget over the support of π.

Following the framework of distributionally robust optimization [Kuhn et al., 2025], we formalize this restriction by introducing a Kullback–Leibler ambiguity set $\mathcal{Q}_\rho$.

Definition 2 (Ambiguity set)**.** *Let $\rho > 0$ be a misspecification budget. The set $\mathcal{Q}_\rho(\pi)$ contains all distributions $q : \Xi \to \mathcal{P}(\Theta \times \mathcal{X})$ within an average Kullback–Leibler neighborhood of $p(\theta, x \mid \xi)$:*

$$\mathcal{Q}_\rho(\pi) \coloneqq \left\{ q(\theta, x \mid \xi) \;\middle|\; \mathbb{E}_\pi \left[\mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \,\big\|\, p(\theta, x \mid \xi) \big] \right] \le \rho \right\},$$

where $p(\theta, x \mid \xi)$ denotes the experimenter's nominal model.

This global constraint allows nature to strategically allocate its budget under a bounded average distortion, attacking informative design regions more aggressively while ignoring uninformative ones.

Robust Expected Information Gain. We study the inner minimization problem $\inf_{q\in\mathcal{Q}_\rho} V(\pi, q)$, which defines the worst-case utility from the experimenter's perspective. By Lagrangian dualization [Boyd and Vandenberghe, 2004], the global minimization decomposes into pointwise regularized problems for every design $\xi \in \Xi$, coupled through a shared misspecification budget.

Lemma 1 (Minimization decomposition)**.** *The worst-case criterion reduces to a dual form* $\inf_{q\in\mathcal{Q}_\rho} V(\pi, q) = \sup_{\beta>0} \left\{\mathbb{E}_\pi\left[\mathcal{J}_\beta(\xi)\right] - \beta\rho\right\}$*, where* $\mathcal{J}_\beta(\xi)$ *is the regularized utility:*

$$\mathcal{J}_\beta(\xi) \coloneqq \inf_q \left\{U(\xi, q) + \beta\, \mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \;\big\|\; p(\theta, x \mid \xi)\big]\right\}.$$

The radius ρ, or equivalently its dual parameter β, is a robustness knob. Smaller ρ reflects stronger trust in the nominal model $p(\theta, x \mid \xi)$, while larger ρ permits more perturbations. The next result gives the corresponding pointwise objective in closed form.

Proposition 1 (Robust expected information gain)**.** *For any design* $\xi \in \Xi$ *and regularization parameter* $\beta > 0$*, let* $\alpha \coloneqq \beta/(1+\beta) \in (0, 1)$*. The utility* $\mathcal{J}_\beta(\xi)$ *admits the closed-form expression:*

$$\mathcal{J}_\beta(\xi) = \inf_\nu \mathbb{D}_\alpha\big[p(\theta, x \mid \xi) \;\big\|\; p(\theta)\,\nu(x \mid \xi)\big] = \mathbb{D}_\alpha\big[p(\theta, x \mid \xi) \;\big\|\; p(\theta)\,p_\alpha(x \mid \xi)\big] \coloneqq I^S_\alpha(\theta; x)(\xi), \quad (3)$$

where $\mathbb{D}_\alpha$ *denotes Rényi's divergence of order* α *and the tilted marginal* $p_\alpha(x \mid \xi)$ *is given by:*

$$p_\alpha(x \mid \xi) \propto \left[\mathbb{E}_{p(\theta)}\Big[p(x \mid \theta, \xi)^\alpha\Big]\right]^{1/\alpha}. \quad (4)$$

The quantity $I^S_\alpha(\theta; x)$ is Sibson's α-mutual information [Sibson, 1969, Verdú, 2015, Esposito et al., 2022, 2024], interpreted as a *robust expected information gain* with a misspecification order α.

Remark 1. *Unlike Shannon's mutual information, Sibson's* α*-mutual information is not a symmetric measure of information,* $I^S_\alpha(\theta; x)(\xi) \neq I^S_\alpha(x; \theta)(\xi)$ *[Esposito et al., 2024].*

Corollary 1 (Worst-case generative process)**.** *Let* $\alpha \in (0, 1)$ *be the misspecification order determined by the ambiguity radius* ρ*. The worst-case joint distribution* $q^\star(\theta, x \mid \xi)$ *that minimizes the objective in Proposition 1 is given by the geometric mixture:*

$$q^\star(\theta, x \mid \xi) \propto \Big[p(\theta)\,p_\alpha(x \mid \xi)\Big]^{1-\alpha}\Big[p(\theta, x \mid \xi)\Big]^{\alpha},$$

where the α*-marginal* $p_\alpha(x \mid \xi)$ *is defined by* (4)*. Consequently, the worst-case posterior is*

$$q^\star(\theta \mid x, \xi) \propto p(\theta)\,p(x \mid \theta, \xi)^\alpha. \quad (5)$$

Corollary 1 shows how Sibson's α-mutual information leads to robustness at the Bayesian inference level. The α-tilted posterior (5) arises as the optimal belief update under the worst-case generative process. Therefore, acting consistently with this adversarial assumption requires the experimenter to update beliefs about θ using this tempered rule rather than the standard Bayesian posterior. By down-weighting the misspecified likelihood with $\alpha \in (0, 1)$, this result aligns with generalized Bayesian inference techniques for handling model misspecification [see, e.g., Grünwald, 2012, Watson and Holmes, 2016, Grünwald and van Ommen, 2017, Knoblauch et al., 2022, and references therein].

At the design level, robustness is governed by the monotonicity of Sibson's α-mutual information in α [Esposito et al., 2024]. As $\alpha \to 1$, the ambiguity set contracts, adversarial influence vanishes, and $I^S_\alpha(\theta; x)$ recovers the standard expected information gain in Definition 1. As α decreases the objective becomes more conservative, favoring designs whose information gain is stable under misspecification. In the limit $\alpha \to 0$, nature exerts maximal influence, nullifying any gain.

Proposition 2 (Uninformative design)**.** *Under Assumption 1, as* $\alpha \to 0$*, the robust expected information gain converges to* 0*:*

$$\lim_{\alpha\to 0} \sup_{\xi\in\Xi} I^S_\alpha(\theta; x)(\xi) = \lim_{\alpha\to 0} \sup_{\xi\in\Xi} \mathbb{D}_\alpha\big[p(\theta, x \mid \xi) \;\big\|\; p(\theta)\,p_\alpha(x \mid \xi)\big] = 0.$$

The choice of α is therefore a prior modeling choice, much like the choice of $p(\theta)$. Before observing data from the true data-generating process, the experimenter must specify a nominal model $p(\theta, x \mid \xi)$, based on domain knowledge, and then decide how much trust to place in that model. Values $\alpha \approx 1$ express high confidence and recover behavior close to the standard well-specified setting, while values $\alpha \ll 1$ reflect substantial model uncertainty and favor designs that remain informative under stronger adversarial misspecification. Thus, α should be interpreted as a *subjective* modeling variable that controls the experimenter's assumptions about the set of plausible data-generating processes.

Risk-Sensitive Interpretation. Another interpretation of the robust measure $I^S_\alpha(\theta; x)(\xi)$ from Proposition 1 is given by the following decomposition.

Proposition 3 (Risk-sensitive form)**.** *The robust objective admits the risk-sensitive representation*

$$I^S_\alpha(\theta; x)(\xi) = \frac{\alpha}{\alpha - 1} \log \mathbb{E}_{p(x|\xi)} \left[\exp \left\{ \frac{\alpha - 1}{\alpha} G_\alpha(x, \xi) \right\} \right], \tag{6}$$

where $G_\alpha(x, \xi) \coloneqq \mathbb{D}_\alpha\big[p(\theta \mid x, \xi) \,\big\|\, p(\theta)\big]$ is the robust conditional information gain for a specific outcome-design pair (x, ξ). Additionally, for α in a neighborhood of 1, this representation admits the following second-order expansion

$$I^S_\alpha(\theta; x)(\xi) \approx I(\xi) + \frac{\alpha - 1}{2} \mathbb{V}_{p(\theta, x|\xi)} \left[\log \frac{p(\theta \mid x, \xi)}{p(\theta)} \right]. \tag{7}$$

The form in (6) directly mirrors the interpretation of Shannon's mutual information, defined in (2a), as an expected divergence between prior and posterior in the well-specified setting. However, it identifies the conditional utility $G_\alpha(x, \xi)$ as the Rényi divergence between the nominal posterior and the prior, while the overall robust expected information gain $I^S_\alpha(\theta; x)(\xi)$ aggregates these conditional gains through a risk-sensitive exponential average.

Moreover, the expansion in (7) makes risk-sensitivity explicit. Approximately, $I^S_\alpha(\theta; x)(\xi)$ equals Shannon's expected information gain $I(\xi)$ with an additional variance term that, for $\alpha \in (0, 1)$, penalizes designs whose realized information gain fluctuates across design-outcome pairs.

4 Nested Monte Carlo Estimator

The robust expected information gain $I^S_\alpha \coloneqq I^S_\alpha(\theta; x)$ only has closed-form solutions in special cases. We propose to use a simple nested Monte Carlo estimator to estimate it in the general case. We use the following decomposition of Sibson's α-mutual information for estimation.

Corollary 2 (Nested representation of I^S_α)**.** *The robust expected information gain $I^S_\alpha(\xi)$ can be expressed in terms of nested expectations over nonlinear maps:*

$$I^S_\alpha(\xi) = \frac{\alpha}{\alpha - 1} \log \mathbb{E}_{p(x|\xi)} \left[\left(\mathbb{E}_{p(\theta)} \left[w(x, \theta, \xi)^\alpha \right] \right)^{1/\alpha} \right], \text{ with } w(x, \theta, \xi) \coloneqq p(x \mid \theta, \xi)/p(x \mid \xi).$$

To simplify the exposition, we define the outer logarithmic transformation $f(y) \coloneqq \alpha/(\alpha - 1) \log(y)$ and the inner power function $h(u) \coloneqq u^{1/\alpha}$. Using this notation, the robust objective can be written compactly as $I^S_\alpha(\xi) = f(g(\xi))$, where $g(\xi) \coloneqq \mathbb{E}_{p(x|\xi)} \Big[h\big(\mathbb{E}_{p(\theta)} \left[w(x, \theta, \xi)^\alpha \right] \big) \Big]$.

Definition 3 (Nested Monte Carlo estimator $\tilde{I}^S_\alpha$)**.** *For a design ξ and misspecification order α, the empirical estimator $\tilde{I}^S_\alpha(\xi) \coloneqq f(\tilde{g}(\xi))$ is constructed using a nested Monte Carlo scheme. Let $\{(\theta^{(i)}, x^{(i)})\}_{i=1}^N$ be a set of samples drawn from the joint distribution, and $\{\theta^{(i,j)}\}_{j=1}^M$ an independent set of samples drawn from the prior. We define $\tilde{g}(\xi) \coloneqq \frac{1}{N} \sum_{i=1}^N h\big(\frac{1}{M} \sum_{j=1}^M w(x^{(i)}, \theta^{(i,j)}, \xi)^\alpha \big)$. When the density ratio is intractable, we rely on a contrastive estimator $\tilde{w}$ using K auxiliary samples: $\tilde{w}(x^{(i)}, \theta^{(i,j)}, \xi) \coloneqq p(x^{(i)} \mid \theta^{(i,j)}, \xi) / \frac{1}{K} \sum_{k=1}^K p(x^{(i)} \mid \theta^{(i,k)}, \xi)$.*

Appendix C provides a numerically stable method for computing the estimator. Because the estimator applies the nonlinear functions to empirical averages rather than exact expectations, Jensen's inequality implies that it is biased for any finite sample sizes N and M. Nonetheless, the analysis of Rainforth et al. [2018] establishes that it remains consistent and its mean squared error decays as $\mathcal{O}(1/N + 1/M)$, when w is available in closed form.

Next, we provide explicit bounds for its bias under Assumption 2, which defines regularity constants L_f, L_h, C_h, σ_h, and σ_w for the functions f, h, and w.

Proposition 4 (Bias bound for $\tilde{I}^S_\alpha$)**.** *Let $\tilde{I}^S_\alpha(\xi)$ be the estimator of $I^S_\alpha(\xi)$ given by Definition 3. Under Assumption 2, for any design $\xi \in \Xi$, the bias of $\tilde{I}^S_\alpha(\xi)$ is bounded by:*

$$\left| \mathbb{E}\big[\tilde{I}^S_\alpha(\xi)\big] - I^S_\alpha(\xi) \right| \leq L_f \left(\frac{L_h \sigma_w}{\sqrt{M}} + \sqrt{\frac{\sigma_h^2}{N} + \frac{4 C_h L_h \sigma_w}{N \sqrt{M}}} \right).$$

5 PAC-Bayesian Design Policies

In the previous section, we introduced a nested Monte Carlo estimator $\tilde{I}_\alpha^S(\xi)$ for the robust expected information gain $I_\alpha^S(\xi)$. However, this estimator is a biased and noisy oracle, introducing errors that can mislead design selection if optimized naively. To address this, we adopt stochastic design policies and control the estimator error through a PAC-Bayesian approach. In Appendix D, we provide a discussion of classical PAC analysis and the PAC-Bayesian approach.

We view design selection as a PAC-Bayes continuous bandit problem [Flynn et al., 2023], where the designs ξ are arms and the true rewards are determined by I_α^S, observed only through the noisy surrogate $\tilde{I}_\alpha^S$. The PAC-Bayesian formulation optimizes a high-probability lower bound on the true robust objective by penalizing the divergence to a prior policy π_0, thereby balancing empirical performance against estimator uncertainty [Alquier, 2024].

To establish the PAC-Bayesian guarantee, we leverage results of prior sections to derive a probability bound on the bias of the robust expected information gain estimator.

Proposition 5 (Uniform convergence of $\tilde{I}_\alpha^S$)**.** *Let $\tilde{I}_\alpha^S(\xi)$ be the estimator of $I_\alpha^S(\xi)$ from Definition 3. Under Assumption 2, for any tolerance $t > 0$, and for a sufficiently large inner sample size $M \geq 4L_f^2 L_h^2 \sigma_w^2 / t^2$, the probability that the estimator deviates from the oracle by more than t decays exponentially with the outer sample size N:*

$$\mathbb{P}\left(\left|\tilde{I}_\alpha^S(\xi) - I_\alpha^S(\xi)\right| \geq t\right) \leq 2\exp\left\{-Nt^2/(2L_f^2 C_h^2)\right\}.$$

Proposition 5 effectively states that the absolute error of our naive estimator $\tilde{I}_\alpha^S$ behaves like a sub-Gaussian random variable *if* we pay the cost M to make it so. With this concentration behavior in mind, we can now establish a PAC-Bayes lower bound for the robust expected information gain.

Proposition 6 (PAC-Bayes lower bound for I_α^S)**.** *Let $\pi_0 \in \Pi$ be a prior design policy, $\delta \in (0,1)$ a confidence level, and $\lambda > 0$ a precision parameter. Provided $M \geq 2NL_h^2\sigma_w^2/(C_h^2 \log(2/\delta))$, then with probability at least $1-\delta$, this bound holds for all $\pi \in \Pi$ simultaneously:*

$$\mathbb{E}_\pi\left[I_\alpha^S(\xi)\right] \geq \mathbb{E}_\pi\left[\tilde{I}_\alpha^S(\xi)\right] - \frac{\lambda L_f^2 C_h^2}{4N} - \frac{\mathbb{D}_{\mathrm{KL}}[\pi \parallel \pi_0] + 2\log(2/\delta)}{\lambda}.$$

Proposition 6 yields a high-probability lower bound on the true *robust* objective expressed in terms of the Monte Carlo surrogate and a KL penalty. Since the remaining terms in the bound are independent of π, maximizing this lower bound over Π reduces to the following variational optimization problem:

$$\pi^\star := \arg\max_{\pi \in \Pi} \left\{\mathbb{E}_\pi\left[\tilde{I}_\alpha^S(\xi)\right] - \frac{\mathbb{D}_{\mathrm{KL}}[\pi \parallel \pi_0]}{\lambda}\right\}, \tag{8}$$

whose maximizer is given by a Gibbs policy [Alquier, 2024]: $\pi^\star(\xi) \propto \pi_0(\xi)\,\exp\left\{\lambda\,\tilde{I}_\alpha^S(\xi)\right\}$. This objective can be optimized using stochastic policy search methods, including natural gradient and mirror descent algorithms [Amari, 1998, Kakade, 2001, Beck and Teboulle, 2003, Peters et al., 2010]. The role of the precision parameter λ is discussed in Appendix E.

6 Related Work

The two works most closely related to ours methodologically are Go and Isaac [2022] and Waite and Woods [2022]. Go and Isaac [2022] also adopt a distributionally robust optimization framework, but focus on prior misspecification, leading to a risk-sensitive objective based on log-exponential averages of Kullback–Leibler divergences rather than a Rényi-type divergence. By contrast, our maximin formulation naturally recovers Sibson's α-mutual information. Waite and Woods [2022] study maximin experimental design from a *frequentist* perspective and consider stochastic policies. However, the absence of prior beliefs over θ leads to a different mechanism for constructing robust designs, and their theoretical analysis is confined to linear models.

Two recent works applied generalized posteriors to Bayesian experimental design: Overstall et al. [2025] and Barlas et al. [2025]. In both cases, a Gibbs posterior is introduced as an *ad hoc* device to robustify the inference procedure given a design and a measured outcome. In contrast, in our framework the tilted posterior (5) emerges naturally from the maximin robustness principle as the

experimenter's robust belief update rule. This construction leads to a unified formulation in which misspecification is addressed jointly at the level of posterior inference and the data-generating process. In Appendix G, we discuss the connection to Barlas et al. [2025] in further detail through two concrete experimental design problems, and show how the resulting objectives differ in general.

A related line of work develops alternative utility functions that incorporate information from the *true, but unknown* data-generating process. Catanach and Das [2023] introduce the expected generalized information gain, replacing the KL divergence in (2a) with a discrepancy measure defined under the true process. Since this quantity is not directly computable, their approach requires specifying a model class containing the true process and evaluating the sensitivity of the utility with respect to parameters indexing this class. Similarly, Tang et al. [2025] augment the EIG with a penalty on covariate shifts between the training data collected using optimized designs and a reference test set. Therefore, the proposed acquisition function is constrained by the availability of such test data.

Finally, some complementary approaches address misspecification by augmenting the model with expressive components such as Gaussian processes [Feng, 2015, Forster et al., 2025].

7 Numerical Evaluation

Our evaluation focuses on two aspects. First, we highlight the properties of Sibson's α-MI as a criterion for robust experimental design, emphasizing that robustness manifests at the levels of the posterior update and design selection. Second, when the robust EIG is only available through a biased and noisy empirical estimator, we show that naive deterministic design optimizers can yield suboptimal designs, hence validating the need for the PAC-Bayes policy for design optimization.

To study the properties of Sibson's α-MI, we focus on problem settings in which relevant quantities admit closed-form expressions. This lets us isolate the behavior of Sibson's α-MI without confounding numerical artifacts. Specifically, we consider a continuous linear regression problem and a discrete A/B testing problem, which allow for exact computation of posterior updates, conditional information gains, and Sibson's α-MI. Details on computations are provided in Appendices F.1 and F.2:

1. In linear regression, the experimenter assumes that measurements are real-valued responses $x \in \mathbb{R}$ generated according to the Gaussian $p(x \mid \theta, \xi) = \mathrm{N}(x; \xi\,\theta_1 + \theta_2,\ \sigma^2)$, where $\xi \in [-1, 1]$ is the design and the parameters are $\theta = (\theta_1, \theta_2)^\top \in \mathbb{R}^2$, with prior $p(\theta) = \mathrm{N}(\theta; \mu_0, \Sigma_0)$. The true data-generating process draws from a Student-t distribution $\mathrm{T}(x; \xi\,\theta_1 + \theta_2, \sigma^2, \nu)$ with ν degrees of freedom and the same conditional mean, with heavier tails than the assumed Gaussian.
2. In A/B testing, the experimenter assumes measurements are counts $x = (x_a, x_b) \in \{0, \ldots, n_a\} \times \{0, \ldots, n_b\}$ from a Binomial $p(x \mid \theta, \xi) = \prod_k \mathrm{Bin}(x_k; n_k, \theta_k)$, where $k \in \{a, b\}$. The design $\xi = n_a \in \{0, \ldots, N_x\}$ controls group allocation (with $n_b = N_x - n_a$), and the parameters are $\theta = (\theta_a, \theta_b)^\top \in [0, 1]^2$ with independent Beta priors $p(\theta) = \prod_{k \in \{a,b\}} \mathrm{Beta}(\theta_k; \delta_k, \gamma_k)$. The true data-generating process draws θ from the Beta prior, then for each variant samples a latent rate $r_k \sim \mathrm{Beta}(\kappa\,\theta_k, \kappa\,(1 - \theta_k))$ and measurements $x_k \sim \mathrm{Bin}(n_k, r_k)$, yielding overdispersed Beta-Binomial measurements.

Sibson's α-Mutual Information. We compare the *realized* information gain under nominal and robust design criteria. In the nominal setting, the experimenter quantifies information via the Kullback–Leibler divergence in (2a), whereas the robust formulation replaces this measure with a Rényi divergence, as in (6). Figure 1 reports histograms over 10^4 simulated experiments for both the linear regression and A/B testing settings under uniformly distributed designs. Superficially, the nominal criterion appears to deliver larger gains. This comparison is, however, inherently self-referential. The nominal utility is a *subjective* measure that assumes the nominal likelihood is the true data-generating process, leading the experimenter to overstate the perceived informativeness, whereas the robust utility explicitly accounts for misspecification and yields a more conservative estimate.

This overconfidence manifests at two distinct levels. *At the inference stage*, the coverage analysis in the right panel of Figure 1 compares expected credible levels with the empirical frequency at which the true parameter falls within the corresponding credible sets. For the nominal Bayesian posterior ($\alpha = 1$), coverage curves lie well below the diagonal, indicating severe undercoverage. The robust tilted posterior progressively improves calibration as α decreases, with curves moving closer to or above the diagonal reflecting a more conservative stance compared to the nominal Bayes posterior.

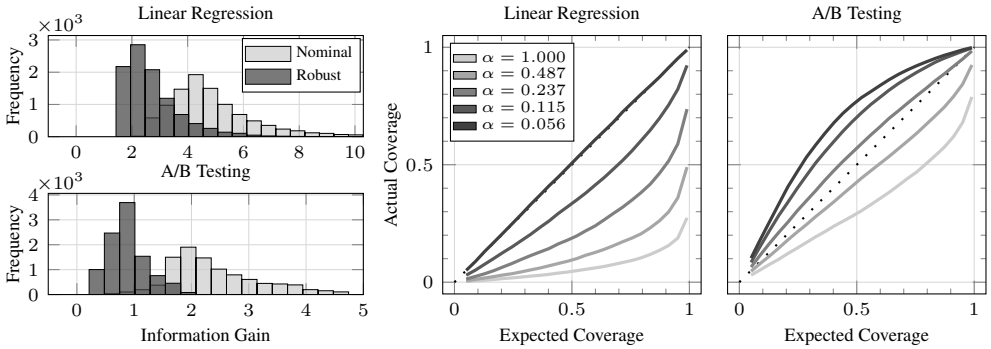

Figure 1: Comparison of nominal and robust formulations for linear regression and A/B testing. *Left:* histograms of realized information gains under nominal ($\alpha = 1.0$) and robust ($\alpha = 0.056$) assumptions. *Right:* expected vs. actual coverage for varying values of α. Nominal posteriors are clearly overconfident; calibration improves for values $\alpha < 1$. Data obtained from 10^4 experiments.

	Linear Regression		A/B Testing	
α	Random	Optimal	Random	Optimal
0.056	4.62×10^{-1}	$\mathbf{2.47 \times 10^{-1}}$	1.01×10^{-1}	$\mathbf{7.49 \times 10^{-2}}$
0.115	4.78×10^{-1}	$\mathbf{2.50 \times 10^{-1}}$	9.03×10^{-2}	$\mathbf{7.13 \times 10^{-2}}$
0.237	5.22×10^{-1}	$\mathbf{2.53 \times 10^{-1}}$	8.98×10^{-2}	$\mathbf{7.31 \times 10^{-2}}$
0.487	5.91×10^{-1}	$\mathbf{2.56 \times 10^{-1}}$	9.77×10^{-2}	$\mathbf{8.88 \times 10^{-2}}$
1.000	6.82×10^{-1}	$\mathbf{2.57 \times 10^{-1}}$	1.09×10^{-1}	$\mathbf{1.06 \times 10^{-1}}$

Table 1: Posterior-mean RMSE for linear regression and A/B testing. We compare the robust posterior under random and optimal designs for varying α. Robust designs that optimize Sibson's MI lead to smaller mean squared errors. Data obtained over 10^4 experiments.

	Design		Sibson RMSE			GEIG RMSE		
α	Sibson	GEIG	Mean	P_{10}	P_{90}	Mean	P_{10}	P_{90}
0.011	0	30	**0.120**	0.034	0.208	0.132	0.036	0.231
0.022	0	33	**0.098**	0.030	0.167	0.112	0.033	0.193
0.041	0	35	**0.081**	0.025	0.143	0.093	0.028	0.158
0.077	0	37	**0.072**	0.022	0.141	0.079	0.025	0.146
0.147	0	39	**0.071**	0.020	0.144	0.076	0.024	0.145
0.278	18	42	**0.075**	0.022	0.150	0.083	0.028	0.154
0.527	33	43	**0.091**	0.030	0.163	0.095	0.031	0.169
1.000	41	42	0.106	0.036	0.184	0.106	0.037	0.184

Table 2: Comparison of Sibson's and GEIG's optimal designs and RMSE on an A/B testing problem with $N_x = 100$. The two objectives lead to different design choices. We report the optimal allocation, mean RMSE, the 10th and 90th percentiles for each method. Evaluation performed over 10^4 experiments.

At design selection, Table 1 reports the root mean squared error (RMSE) of the posterior mean under random and optimal designs across a geometrically spaced grid of α values. Optimal designs produce posterior means that are closer to the true parameter. In both the linear regression and A/B testing settings, there exists an $\alpha < 1$ for which the robust optimal design outperforms the nominal optimum at $\alpha = 1$. Overall, these results confirm that Sibson's α-mutual information favors designs that are less sensitive to model misspecification and leads to better calibrated posteriors.

Additionally, in Appendix G, we give a detailed comparison between Sibson's α-mutual information and the *Gibbs expected information gain* (GEIG) proposed by Barlas et al. [2025]. Although both objectives are connected to tilted Gibbs posteriors, we show that they are not equivalent in general. Table 2 includes an empirical comparison across α values, demonstrating that the two objectives can lead to different optimal designs on the A/B testing problem.

PAC-Bayes Policies. While the preceding evaluation relied on closed-form formulas for computation, this section focuses on optimizing designs using the nested Monte Carlo estimator, thereby foregoing any benefits of analytical tractability. To test design optimization in a higher-dimensional setting, we extend the linear regression problem to a 10-dimensional parameter and design space, while A/B testing accommodates $N_x = 100$ subjects. In both cases, the prior policy π_0 is taken to be uniform over the design space and the precision is tuned over a grid and set to $\lambda = 10^6$; see Table 5.

Figure 2 depicts the failure of naive design optimization when the objective is estimated by a noisy, biased oracle. We compare stochastic gradient descent for linear regression and exhaustive enumeration for A/B testing against a PAC-Bayes stochastic policy optimized via mirror descent, while fixing α, λ, and the sample sizes N and M. We initialize 1024 distinct naive optimizers and evaluate the realized regret and design optimality relative to the tractable optimum. The histograms

Table 3: Regret comparison on a linear regression between a naive optimizer and a PAC-Bayes policy for $M = 16$ and variable N. We report the mean regret together with the 10th and 90th percentiles, computed over 256 repetitions.

N	Naive			PAC-Bayes		
	Mean	P_{10}	P_{90}	Mean	P_{10}	P_{90}
16	0.327	0.219	0.430	**0.077**	**0.058**	**0.099**
32	0.312	0.212	0.413	**0.083**	**0.040**	**0.135**
64	0.307	0.205	0.407	**0.042**	**0.031**	**0.054**
128	0.296	0.200	0.386	**0.041**	**0.036**	**0.046**
256	0.282	0.190	0.370	**0.018**	**0.013**	**0.023**

Table 4: Relative regret comparison on a linear regression between a naive optimizer and a PAC-Bayes policy for varying α values. We report the mean relative regret together with the 10th and 90th percentiles, computed over 256 repetitions.

α	Naive			PAC-Bayes		
	Mean	P_{10}	P_{90}	Mean	P_{10}	P_{90}
0.056	0.098	0.066	0.128	**0.023**	**0.014**	**0.031**
0.115	0.100	0.066	0.132	**0.020**	**0.013**	**0.028**
0.237	0.103	0.072	0.138	**0.016**	**0.012**	**0.021**
0.487	0.104	0.070	0.143	**0.025**	**0.016**	**0.035**
1.000	0.105	0.066	0.147	**0.025**	**0.016**	**0.036**

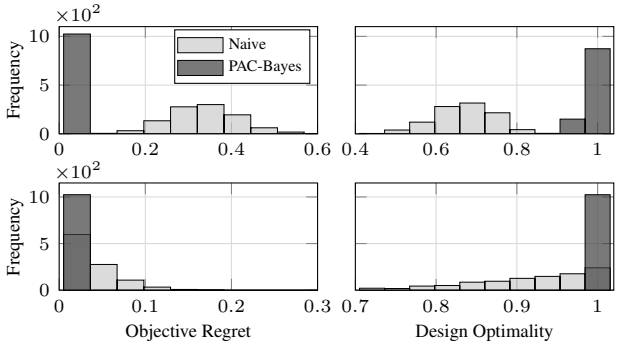

Figure 2: Empirical distributions of regret (left) and design optimality (right) for linear regression (top) and A/B testing (bottom). We compare naive optimization against our PAC-Bayes policy. Naive optimization has higher regret and variability. Its designs are suboptimal, reflected in lower similarity scores relative to the theoretical optimal design. Data from 1024 trials.

show that naive optimizers incur higher and more variable regret, reflecting suboptimal designs, whereas the PAC-Bayes policy concentrates near the optimum and achieves consistently lower regret.

Finally, we examine the impact of sample complexity in Table 3, which reports the mean regret alongside the 10th and 90th percentiles across varying outer sample sizes N. While the performance of the naive optimizer marginally improves as N increases, it fails to match the performance of the PAC-Bayes policy. Table 4 confirms that this performance gap holds uniformly across different values of α. For fixed sample sizes N and M, the naive optimizer consistently incurs higher regret than the PAC-Bayes policy, validating the PAC-Bayesian perspective.

8 Discussion

We established a principled framework for robust Bayesian experimental design using the maximin principle. We proposed Sibson's α-mutual information as a robust alternative to the standard expected information gain and showed that the resulting belief update is an α-tilted posterior, while the conditional information gain corresponds to a Rényi divergence between posterior and prior. To handle the intractability of Sibson's α-MI and the finite-sample Monte Carlo error, we introduced PAC-Bayesian stochastic policies and proved a lower bound for the true robust objective. Together, these results characterize maximin robust Bayesian experimental design under information-theoretic constraints, including the adversarial perturbation, robust information gain, and tilted posterior update.

As for limitations, although the maximin formulation provides worst-case guarantees, it may lead to overly conservative designs [Watson and Holmes, 2016]. Exploring alternative robust decision-making paradigms could complement our work. In addition, calibrating the misspecification order α as experimental data accumulates remains an open challenge. Related calibration approaches in the generalized Bayesian inference literature may offer useful intuition [Wu and Martin, 2021]. A further research direction is robust *sequential* experimental design, where the experimenter plans over a horizon of experiments rather than myopically optimizing a single design [Foster et al., 2021].

Acknowledgments and Disclosure of Funding

H. Abdulsamad and C. A. Naesseth acknowledge the support of the UvA–Bosch Delta Lab and the Amsterdam Machine Learning Lab (AMLab) at the University of Amsterdam. This project was supported by the Bosch Center for Artificial Intelligence. The authors declare no competing interests.

References

Pierre Alquier. User-friendly introduction to PAC-Bayes bounds. *Foundations and Trends® in Machine Learning*, 17(2):174–303, 2024.

Shun-Ichi Amari. Natural gradient works efficiently in learning. *Neural Computation*, 10(2):251–276, 1998.

Yasir Zubayr Barlas, Sabina J Sloman, and Samuel Kaski. Robust experimental design via generalised Bayesian inference. *arXiv preprint arXiv:2511.07671v1*, 2025.

Tamer Başar and Geert Jan Olsder. *Dynamic Noncooperative Game Theory*. SIAM, 1998.

Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, 2003.

James O. Berger. *Statistical Decision Theory and Bayesian Analysis*. Springer New York, 2nd edition, 1985.

José M. Bernardo. Expected information as expected utility. *The Annals of Statistics*, 7(3):686–690, 1979.

José M. Bernardo and Adrian F. M. Smith. *Bayesian Theory*. Wiley, 1994.

Pier Giovanni Bissiri, Chris C. Holmes, and Stephen G. Walker. A general framework for updating belief distributions. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 78 (5):1103–1130, 2016.

Tom Blau, Edwin V. Bonilla, Iadine Chades, and Amir Dezfouli. Optimizing sequential experimental design with deep reinforcement learning. In *International Conference on Machine Learning*, 2022.

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

Tommie A. Catanach and Niladri Das. Metrics for Bayesian optimal experiment design under model misspecification. In *IEEE Conference on Decision and Control*, pages 7707–7714, 2023.

Olivier Catoni. PAC-Bayesian supervised classification: The thermodynamics of statistical learning. *Institute of Mathematical Statistics Lecture Notes - Monograph Series*, 56:1–163, 2007.

Kathryn Chaloner and Isabella Verdinelli. Bayesian experimental design: A review. *Statistical Science*, 10(3):273–304, 1995.

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

Monroe D Donsker and SR Srinivasa Varadhan. Asymptotic evaluation of certain Markov process expectations for large time. *Communications on Pure and Applied Mathematics*, 28(1):1–47, 1975.

Amedeo Roberto Esposito, Adrien Vandenbroucque, and Michael Gastpar. On Sibson's α-mutual information. In *IEEE International Symposium on Information Theory*, pages 2904–2909, 2022.

Amedeo Roberto Esposito, Michael Gastpar, and Ibrahim Issa. Sibson's α-mutual information and its variational representations. *arXiv preprint arXiv:2405.08352*, 2024.

Chi Feng. Optimal Bayesian experimental design in the presence of model error. Master's thesis, Massachusetts Institute of Technology, 2015.

Hamish Flynn, David Reeb, Melih Kandemir, and Jan Peters. PAC-Bayes bounds for bandit problems: A survey and experimental comparison. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(12):15308–15327, 2023.

Alex Forster, Desi R. Ivanova, and Tom Rainforth. Improving robustness to model misspecification in Bayesian experimental design. In *Symposium of Advances of Approximate Bayesian Inference*, 2025.

Adam Foster, Desi R Ivanova, Ilyas Malik, and Tom Rainforth. Deep adaptive design: Amortizing sequential Bayesian experimental design. In *International Conference on Machine Learning*, 2021.

Abhik Ghosh and Ayanendranath Basu. Robust Bayes estimation using the density power divergence. *Annals of the Institute of Statistical Mathematics*, 68(2):413–437, 2016.

Tilmann Gneiting and Adrian E Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007.

Jinwoo Go and Tobin Isaac. Robust expected information gain for optimal Bayesian experimental design using ambiguity sets. In *Conference on Uncertainty in Artificial Intelligence*, 2022.

Peter Grünwald. The safe Bayesian: Learning the learning rate via the mixability gap. In *International Conference on Algorithmic Learning Theory*, 2012.

Peter Grünwald and Thijs van Ommen. Inconsistency of Bayesian inference for misspecified linear models, and a proposal for repairing it. *Bayesian Analysis*, 12(4):1069–1103, 2017.

Peter D. Grünwald and A. Philip Dawid. Game theory, maximum entropy, minimum discrepancy and robust Bayesian decision theory. *The Annals of Statistics*, 32(4):1367–1433, 2004.

Tuomas Haarnoja, Aurick Zhou, Kristian Hartikainen, George Tucker, Sehoon Ha, Jie Tan, Vikash Kumar, Henry Zhu, Abhishek Gupta, Pieter Abbeel, et al. Soft actor-critic algorithms and applications. *arXiv preprint arXiv:1812.05905*, 2018.

Chris C. Holmes and Stephen G. Walker. Assigning a value to a power likelihood in a general Bayesian model. *Biometrika*, 104(2):497–503, 2017.

Yifan Hu, Xin Chen, and Niao He. Sample complexity of sample average approximation for conditional stochastic optimization. *SIAM Journal on Optimization*, 30(3):2103–2133, 2020.

Sahel Iqbal, Adrien Corenflos, Simo Särkkä, and Hany Abdulsamad. Nesting particle filters for experimental design in dynamical systems. In *International Conference on Machine Learning*, 2024.

Sham M Kakade. A natural policy gradient. *Advances in neural information processing systems*, 14, 2001.

Jeremias Knoblauch, Jack Jewson, and Theodoros Damoulas. An optimization-centric view on Bayes' rule: Reviewing and generalizing variational inference. *Journal of Machine Learning Research*, 23 (132):1–109, 2022.

Daniel Kuhn, Soroosh Shafiee, and Wolfram Wiesemann. Distributionally robust optimization. *Acta Numerica*, 34:579–804, 2025.

Juliane Liepe, Sarah Filippi, Michał Komorowski, and Michael PH Stumpf. Maximizing the information content of experiments in systems biology. *PLoS Computational Biology*, 9(1), 2013.

Dennis Victor Lindley. On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, 27(4):986–1005, 1956.

David A McAllester. Some PAC-Bayesian theorems. In *Conference on Computational Learning Theory*, pages 230–234, 1998.

JA Melendez, RJ Furnstahl, HW Grießhammer, JA McGovern, DR Phillips, and MT Pratola. Designing optimal experiments: An application to proton Compton scattering. *The European Physical Journal A*, 57(3):81, 2021.

Antony M Overstall, Jacinta Holloway-Brown, and James M McGree. Gibbs optimal design of experiments. *arXiv preprint arXiv:2310.17440*, 2025.

Jan Peters, Katharina Mulling, and Yasemin Altun. Relative entropy policy search. In *AAAI Conference on Artificial Intelligence*, pages 1607–1612, 2010.

Ian R Petersen, Matthew R James, and Paul Dupuis. Minimax optimal control of stochastic uncertain systems with relative entropy constraints. *IEEE Transactions on Automatic Control*, 45(3):398–412, 2000.

Tom Rainforth, Rob Cornish, Hongseok Yang, Andrew Warrington, and Frank Wood. On nesting Monte Carlo estimators. In *International Conference on Machine Learning*, volume 80, pages 4267–4276, 2018.

Tom Rainforth, Adam Foster, Desi R. Ivanova, and Freddie Bickford Smith. Modern Bayesian experimental design. *Statistical Science*, 39(1):100–114, 2024.

Yannick Rondelez. Competition for catalytic resources alters biological network dynamics. *Physical Review Letters*, 108(1):018102, 2012.

Leonard J Savage. Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, 66(336):783–801, 1971.

Matthias Schultheis, Boris Belousov, Hany Abdulsamad, and Jan Peters. Receding horizon curiosity. In *Conference on Robot Learning*, pages 1278–1288. PMLR, 2020.

Ben Shababo, Brooks Paige, Ari Pakman, and Liam Paninski. Bayesian inference and online experimental design for mapping neural microcircuits. In *International Conference on Neural Information Processing Systems*, 2013.

Robin Sibson. Information radius. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 14(2):149–160, 1969.

Roubing Tang, Sabina J Sloman, and Samuel Kaski. Generalization analysis for Bayesian optimal experiment design under model misspecification. *arXiv preprint arXiv:2506.07805*, 2025.

Bob van Sluijs, Tao Zhou, Britta Helwig, Mathieu G Baltussen, Frank HT Nelissen, Hans A Heus, and Wilhelm TS Huck. Iterative design of training data to control intricate enzymatic reaction networks. *Nature Communications*, 15(1):1602, 2024.

Sergio Verdú. α-mutual information. In *2015 Information Theory and Applications Workshop (ITA)*, pages 1–6. IEEE, 2015.

Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*, volume 47. Cambridge university press, 2018.

John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

Martin J. Wainwright. *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*. Cambridge University Press, 2019.

Timothy W. Waite and David C. Woods. Minimax efficient random experimental design strategies with application to model-robust design for prediction. *Journal of the American Statistical Association*, 117(539):1452–1465, 2022.

Abraham Wald. *Statistical Decision Functions*. Wiley, 1950.

James Watson and Chris Holmes. Approximate models and robust decisions. *Statistical Science*, 31 (4):465–489, 2016.

Peter Whittle. *Risk-Sensitive Optimal Control*. Wiley, 1990.

Pei-Shien Wu and Ryan Martin. Calibrating generalized predictive distributions. *arXiv preprint arXiv:2107.01688*, 2021.

Tong Zhang. From ϵ-entropy to KL-entropy: Analysis of minimum information complexity density estimation. *The Annals of Statistics*, 34(5):2180–2210, 2006.

A Organization of the Appendix

The Appendix is organized as follows:

- Appendix B provides proofs for the various lemmata and propositions.
- Appendix C provides details on the nested Monte Carlo estimator.
- Appendix D compares PAC-Bayesian and deterministic policies.
- Appendix E discusses the role and choice of the PAC-Bayes precision.
- Appendix F provides details for the numerical evaluations.
- Appendix G explores the connection to Barlas et al. [2025].

B Proofs

B.1 Proof of Lemma 1

We consider the worst-case expected utility minimization problem defined by:

$$\begin{aligned}\inf_q V(\pi, q) &= \inf_q \mathbb{E}_\pi \left[U(\xi, q)\right] \\ &= \inf_q \mathbb{E}_\pi \left[\mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \;\big\|\; p(\theta)\, q(x \mid \xi)\big]\right],\end{aligned}$$

subject to the Kullback–Leibler constraint averaged under $\pi(\xi)$:

$$\mathbb{E}_\pi \left[\mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \;\big\|\; p(\theta, x \mid \xi)\big]\right] \leq \rho.$$

Given that both the objective and the constraint are convex in $q(\cdot \mid \xi)$, and under strict feasibility assumptions, strong duality holds [Boyd and Vandenberghe, 2004]. Therefore, we can write:

$$\inf_{q \in \mathcal{Q}_\rho} V(\pi, q) = \sup_{\beta > 0} \inf_q \mathcal{L}(\beta, \pi, q),$$

where $\mathcal{L}$ is the Lagrangian functional given by:

$$\mathcal{L}(\beta, \pi, q) \coloneqq \mathbb{E}_\pi \left[\mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \;\big\|\; p(\theta)\, q(x \mid \xi)\big]\right] + \beta\, \mathbb{E}_\pi \left[\mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \;\big\|\; p(\theta, x \mid \xi)\big]\right] - \beta\rho.$$

As a result, the worst-case expected utility is an expectation over regularized pointwise objectives:

$$\inf_{q \in \mathcal{Q}_\rho} V(\pi, q) = \sup_{\beta > 0} \left\{\mathbb{E}_\pi \left[\mathcal{J}_\beta(\xi)\right] - \beta\rho\right\},$$

where $\mathcal{J}_\beta(\xi)$ is the regularized worst-case utility:

$$\mathcal{J}_\beta(\xi) \coloneqq \inf_q \left\{U(\xi, q) + \beta\, \mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \;\big\|\; p(\theta, x \mid \xi)\big]\right\}.$$

□

B.2 Proof of Proposition 1

We start by stating a useful lemma that highlights a variational interpretation of the mutual information.

Lemma 2 (Shannon's variational mutual information, Verdú, 2015)**.** *Shannon's mutual information between θ and x given ξ is the minimal value of the Kullback–Leibler divergence between the true joint distribution and the product of variational marginals:*

$$I(\xi) \coloneqq \inf_{\nu \in \mathcal{P}(\mathcal{X})} \mathbb{D}_{\mathrm{KL}}\big[p(\theta, x \mid \xi) \;\big\|\; p(\theta)\, \nu(x \mid \xi)\big].$$

The infimum is attained when the variational marginal coincides with the true marginal, so that $\nu^\star(x \mid \xi) = p(x \mid \xi)$. At this optimum, we obtain:

$$I(\xi) = \mathbb{D}_{\mathrm{KL}}\big[p(\theta, x \mid \xi) \;\big\|\; p(\theta)\, p(x \mid \xi)\big].$$

Now, we start from the definition of $\mathcal{J}_\beta(\xi)$ for any $\beta > 0$:

$$\mathcal{J}_\beta(\xi) \coloneqq \inf_q \left\{U(\xi, q) + \beta\, \mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \;\big\|\; p(\theta, x \mid \xi)\big]\right\}.$$

The expression for $U(\xi, q)$ contains the predictive marginal $q(x \mid \xi)$, which changes with the optimized joint law $q(\theta, x \mid \xi)$. To externalize this marginal dependence, let us first make use of the chain rule for relative entropy,

$$\begin{aligned} U(\xi, q) &= \mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \,\big\|\, p(\theta)\, q(x \mid \xi)\big] \\ &= \mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \,\big\|\, q(\theta \mid \xi)\, q(x \mid \xi)\big] + \mathbb{D}_{\mathrm{KL}}\big[q(\theta \mid \xi) \,\big\|\, p(\theta)\big], \end{aligned}$$

then apply the variational form from Lemma 2 to the first term:

$$\mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \,\big\|\, q(\theta \mid \xi)\, q(x \mid \xi)\big] = \inf_{\nu} \mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \,\big\|\, q(\theta \mid \xi)\, \nu(x \mid \xi)\big],$$

with $\nu^\star(x \mid \xi) = q(x \mid \xi)$. Therefore, we can write the payoff as follows:

$$\begin{aligned} U(\xi, q) &= \inf_{\nu} \mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \,\big\|\, q(\theta \mid \xi)\, \nu(x \mid \xi)\big] + \mathbb{D}_{\mathrm{KL}}\big[q(\theta \mid \xi) \,\big\|\, p(\theta)\big] \\ &= \inf_{\nu} \mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \,\big\|\, p(\theta)\, \nu(x \mid \xi)\big]. \end{aligned}$$

This trick temporarily decouples the marginal from the optimized joint distribution. Now, we can write the following minimization problem:

$$\mathcal{J}_\beta(\xi) = \inf_{q} \inf_{\nu} \Big\{ \mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \,\big\|\, p(\theta)\, \nu(x \mid \xi)\big] + \beta\, \mathbb{D}_{\mathrm{KL}}\big[q(\theta, x \mid \xi) \,\big\|\, p(\theta, x \mid \xi)\big] \Big\},$$

with implicit constraints to ensure that $q(\theta, x \mid \xi)$ and $\nu(x \mid \xi)$ are proper densities.

Since q and ν are optimized over independent sets, $\mathcal{P}(\Theta \times \mathcal{X})$ and $\mathcal{P}(\mathcal{X})$, the objective may be viewed as a functional on their product space, and the nested minimization commutes to $\inf_\nu \inf_q$. Furthermore, for a fixed marginal ν, the minimization over q is convex with a convex constraint and strong duality holds under strict feasibility assumptions [Boyd and Vandenberghe, 2004].

To form the associated Lagrangian functional, we introduce the multipliers η and κ associated with the normalization constraints. The resulting dual problem is:

$$\mathcal{J}_\beta(\xi) = \sup_{\kappa} \inf_{\nu} \sup_{\eta} \inf_{q} \mathcal{G}(\beta, \eta, q, \kappa, \nu, \xi),$$

where the Lagrangian functional $\mathcal{G}$ is given by:

$$\begin{aligned} \mathcal{G}(\beta, \eta, q, \kappa, \nu, \xi) = &\int q(\theta, x \mid \xi) \log \frac{q(\theta, x \mid \xi)}{p(\theta)\, \nu(x \mid \xi)} \,\mathrm{d}\theta\, \mathrm{d}x \\ &+ \beta \int q(\theta, x \mid \xi) \log \frac{q(\theta, x \mid \xi)}{p(\theta, x \mid \xi)} \,\mathrm{d}\theta\, \mathrm{d}x \\ &+ \eta \left[\int q(\theta, x \mid \xi) \,\mathrm{d}\theta\, \mathrm{d}x - 1 \right] + \kappa \left[\int \nu(x \mid \xi) \,\mathrm{d}x - 1 \right]. \end{aligned} \tag{9}$$

We start by minimizing $\mathcal{G}$ with respect to the joint distribution $q(\cdot \mid \xi)$. We take the first variation of $\mathcal{G}$ and set it to zero to obtain a stationarity condition:

$$(1 + \beta) \big[\log q^\star(\theta, x \mid \xi) + 1\big] - \log p(\theta) - \log \nu(x \mid \xi) - \beta \log p(\theta, x \mid \xi) + \eta = 0.$$

Solving for $q^\star(\cdot \mid \xi)$:

$$q^\star(\theta, x \mid \xi) = \exp\left\{ -\frac{\eta}{1 + \beta} - 1 \right\} \Big[p(\theta)\, \nu(x \mid \xi)\Big]^{1/(1+\beta)} \Big[p(\theta, x \mid \xi)\Big]^{\beta/(1+\beta)}$$

and accounting for the optimal normalization multiplier $\eta^\star$, we get the geometric mixture:

$$q^\star(\theta, x \mid \xi) = \frac{\Big[p(\theta)\, \nu(x \mid \xi)\Big]^{1/(1+\beta)} \Big[p(\theta, x \mid \xi)\Big]^{\beta/(1+\beta)}}{\displaystyle\int \Big[p(\theta)\, \nu(x \mid \xi)\Big]^{1/(1+\beta)} \Big[p(\theta, x \mid \xi)\Big]^{\beta/(1+\beta)} \,\mathrm{d}\theta\, \mathrm{d}x},$$

where $\beta/(1 + \beta)$ interpolates geometrically between the factorized reference and the nominal joint. Substituting the optimum $q^\star(\cdot \mid \xi)$ back into the Lagrangian (9) eliminates the dependence and

reduces the functional to depend only on the variational marginal $\nu(\cdot \mid \xi)$ and the multipliers β and κ:

$$\begin{aligned}
&\mathcal{G}(\beta, \kappa, \nu, \xi) \\
&= -(1+\beta) \log \left\{ \int \Big[p(\theta)\, \nu(x \mid \xi) \Big]^{1/(1+\beta)} \Big[p(\theta, x \mid \xi) \Big]^{\beta/(1+\beta)} \, \mathrm{d}\theta \, \mathrm{d}x \right\} + \kappa \left[\int \nu(x \mid \xi) \, \mathrm{d}x - 1 \right] \\
&= -(1+\beta) \log \left\{ \int \Big[p(\theta) \nu(x \mid \xi) \Big]^{1/(1+\beta)} \Big[p(\theta) p(x \mid \theta, \xi) \Big]^{\beta/(1+\beta)} \, \mathrm{d}\theta \, \mathrm{d}x \right\} + \kappa \left[\int \nu(x \mid \xi) \, \mathrm{d}x - 1 \right] \\
&= -(1+\beta) \log \left\{ \int \Big[\nu(x \mid \xi) \Big]^{1/(1+\beta)} f_\beta(x \mid \xi) \, \mathrm{d}x \right\} + \kappa \left[\int \nu(x \mid \xi) \, \mathrm{d}x - 1 \right],
\end{aligned}$$

where we have introduced a shorthand for the *soft* marginal likelihood:

$$f_\beta(x \mid \xi) = \int p(\theta) \Big[p(x \mid \theta, \xi) \Big]^{\beta/(1+\beta)} \, \mathrm{d}\theta.$$

We now consider the first variation of $\mathcal{G}$ with respect to $\nu(\cdot \mid \xi)$, yielding the stationarity condition:

$$\kappa - \frac{f_\beta(x \mid \xi) \Big[\nu^\star(x \mid \xi) \Big]^{-\beta/(1+\beta)}}{\int \Big[\nu^\star(x \mid \xi) \Big]^{1/(1+\beta)} f_\beta(x \mid \xi) \, \mathrm{d}x} = 0.$$

Similar to $q(\cdot \mid \xi)$, solving for $\nu^\star(\cdot \mid \xi)$ while accounting for the optimal normalization multiplier $\kappa^\star$ leads to:

$$\nu^\star(x \mid \xi) = \frac{\Big[f_\beta(x \mid \xi) \Big]^{(1+\beta)/\beta}}{\int \Big[f_\beta(x \mid \xi) \Big]^{(1+\beta)/\beta} \mathrm{d}x} = \frac{\left[\int p(\theta) \Big[p(x \mid \theta, \xi) \Big]^{\beta/(1+\beta)} \, \mathrm{d}\theta \right]^{(1+\beta)/\beta}}{\int \left[\int p(\theta) \Big[p(x \mid \theta, \xi) \Big]^{\beta/(1+\beta)} \, \mathrm{d}\theta \right]^{(1+\beta)/\beta} \mathrm{d}x} \coloneqq p_\beta(x \mid \xi).$$

This distribution is a tilted marginal induced by the constraint. Its dependence on β encodes how aggressively mass is concentrated on high-likelihood regions. Substitute this result back into $\mathcal{G}$ to get the dual as a function of β only:

$$\begin{aligned}
\mathcal{G}(\beta, \xi) &= -\beta \, \log \left\{ \int \Big[f_\beta(x \mid \xi) \Big]^{(1+\beta)/\beta} \mathrm{d}x \right\} \\
&= -\beta \, \log \left\{ \int \left[\int p(\theta) \Big[p(x \mid \theta, \xi) \Big]^{\beta/(1+\beta)} \, \mathrm{d}\theta \right]^{(\beta+1)/\beta} \mathrm{d}x \right\} \\
&= -\beta \, \log \left\{ \int \Big[p(\theta, x \mid \xi) \Big]^{\beta/(1+\beta)} \Big[p(\theta)\, p_\beta(x \mid \xi) \Big]^{1/(1+\beta)} \, \mathrm{d}\theta \, \mathrm{d}x \right\} \\
&= \mathbb{D}_{\frac{\beta}{1+\beta}} \big[p(\theta, x \mid \xi) \,\|\, p(\theta)\, p_\beta(x \mid \xi) \big] = \inf_\nu \mathbb{D}_{\frac{\beta}{1+\beta}} \big[p(\theta, x \mid \xi) \,\|\, p(\theta)\, \nu(x \mid \xi) \big],
\end{aligned}$$

where $\mathbb{D}_\zeta$ is the Rényi divergence of order ζ. We substitute $\alpha = \beta/(1+\beta)$ with $\alpha \in (0, 1)$:

$$\mathcal{J}_\beta(\xi) = \mathbb{D}_\alpha \big[p(\theta, x \mid \xi) \;\big\|\; p(\theta)\, p_\alpha(x \mid \xi) \big] \coloneqq I^S_\alpha(\theta; x)(\xi).$$

Finally, the parameter β can be optimized according to the global objective from Lemma 1:

$$\beta^\star \coloneqq \arg\sup_{\beta > 0} \left\{ \mathbb{D}_{\frac{\beta}{1+\beta}} \big[p(\theta, x \mid \xi) \,\|\, p(\theta)\, p_\beta(x \mid \xi) \big] - \beta\, \rho \right\}.$$

□

B.3 Proof of Proposition 3

We start from the definition of the Rényi divergence of order α for distributions P and Q:

$$\mathbb{D}_\alpha[P \parallel Q] = \frac{1}{\alpha - 1} \log \int P(z)^\alpha \, Q(z)^{1-\alpha} \, \mathrm{d}z.$$

For the robust expected information gain (3), we have $P = p(\theta)\, p(x \mid \theta, \xi)$ and $Q = p(\theta)\, p_\alpha(x \mid \xi)$. Substituting leads to:

$$\begin{aligned} I^S_\alpha(\theta; x)(\xi) &= \frac{1}{\alpha - 1} \log \iint \Big[p(\theta)\, p(x \mid \theta, \xi)\Big]^\alpha \Big[p(\theta)\, p_\alpha(x \mid \xi)\Big]^{1-\alpha} \,\mathrm{d}\theta \,\mathrm{d}x \\ &= \frac{1}{\alpha - 1} \log \int p_\alpha(x \mid \xi)^{1-\alpha} \int p(\theta)\, p(x \mid \theta, \xi)^\alpha \,\mathrm{d}\theta \,\mathrm{d}x. \end{aligned}$$

The optimal tilted marginal $p_\alpha(x \mid \xi)$ is defined by the α-tilted marginal:

$$p_\alpha(x \mid \xi) = \frac{1}{Z} \left[\int p(\theta)\, p(x \mid \theta, \xi)^\alpha \,\mathrm{d}\theta\right]^{1/\alpha}, \quad Z = \int \left[\int p(\theta)\, p(x \mid \theta, \xi)^\alpha \,\mathrm{d}\theta\right]^{1/\alpha} \mathrm{d}x.$$

Substituting this form back into the integral, we obtain:

$$\begin{aligned} I^S_\alpha(\theta; x)(\xi) &= \frac{1}{\alpha - 1} \log \left\{ Z^{\alpha-1} \int \left[\int p(\theta)\, p(x \mid \theta, \xi)^\alpha \,\mathrm{d}\theta\right]^{1/\alpha} \mathrm{d}x \right\} \\ &= \frac{\alpha}{\alpha - 1} \log \int \left[\int p(\theta)\, p(x \mid \theta, \xi)^\alpha \,\mathrm{d}\theta\right]^{1/\alpha} \mathrm{d}x. \end{aligned}$$

Next, we multiply and divide the inner term by $p(x \mid \xi)^\alpha$:

$$\begin{aligned} I^S_\alpha(\theta; x)(\xi) &= \frac{\alpha}{\alpha - 1} \log \int p(x \mid \xi) \left[\int p(\theta)\, \frac{p(x \mid \xi, \theta)^\alpha}{p(x \mid \xi)^\alpha} \,\mathrm{d}\theta\right]^{1/\alpha} \mathrm{d}x \\ &= \frac{\alpha}{\alpha - 1} \log \int p(x \mid \xi) \left[\int p(\theta)\, \left[\frac{p(\theta \mid x, \xi)}{p(\theta)}\right]^\alpha \mathrm{d}\theta\right]^{1/\alpha} \mathrm{d}x \\ &= \frac{\alpha}{\alpha - 1} \log \int p(x \mid \xi) \left[\int p(\theta)^{1-\alpha}\, p(\theta \mid x, \xi)^\alpha \,\mathrm{d}\theta\right]^{1/\alpha} \mathrm{d}x. \end{aligned}$$

The inner integral is the scaled exponential of the Rényi divergence between posterior and prior:

$$I^S_\alpha(\theta; x)(\xi) = \frac{\alpha}{\alpha - 1} \log \int p(x \mid \xi) \exp\left\{\frac{\alpha - 1}{\alpha} \mathbb{D}_\alpha\big[p(\theta \mid x, \xi) \,\big\|\, p(\theta)\big]\right\} \mathrm{d}x.$$

To highlight risk-sensitivity, we expand this representation around $\alpha = 1$. Let $\delta \coloneqq \alpha - 1$ and let:

$$G_\alpha(x, \xi) \coloneqq \mathbb{D}_\alpha\big[p(\theta \mid x, \xi) \,\big\|\, p(\theta)\big], \quad L(\theta, x, \xi) \coloneqq \log\, \big[p(\theta \mid x, \xi)/p(\theta)\big]\,.$$

We first expand the conditional Rényi divergence $G_\alpha(x, \xi)$:

$$\begin{aligned} G_\alpha(x, \xi) &= \frac{1}{\alpha - 1} \log \int p(\theta \mid x, \xi)\, \left[\frac{p(\theta \mid x, \xi)}{p(\theta)}\right]^{\alpha-1} \mathrm{d}\theta \\ &= \frac{1}{\delta} \log \mathbb{E}_{p(\theta \mid x, \xi)} \Big[\exp\Big\{\delta\, L(\theta, x, \xi)\Big\}\Big] \\ &= \mathbb{E}_{p(\theta \mid x, \xi)}\, \big[L(\theta, x, \xi)\big] + \frac{\delta}{2}\, \mathbb{V}_{p(\theta \mid x, \xi)} \big[L(\theta, x, \xi)\big] + \mathcal{O}(\delta^2). \end{aligned}$$

Next, we set $\eta \coloneqq (\alpha - 1)/\alpha$. Since $\delta \coloneqq \alpha - 1$, we have:

$$\eta = \frac{\delta}{1 + \delta} = \delta + \mathcal{O}(\delta^2).$$

Expanding the outer exponential average gives:

$$\begin{aligned} I^S_\alpha(\theta; x)(\xi) &= \frac{1}{\eta} \log \mathbb{E}_{p(x \mid \xi)} \Big[\exp\Big\{\eta\, G_\alpha(x, \xi)\Big\}\Big] \\ &= \mathbb{E}_{p(x \mid \xi)}\, \big[G_\alpha(x, \xi)\big] + \frac{\eta}{2}\, \mathbb{V}_{p(x \mid \xi)} \big[G_\alpha(x, \xi)\big] + \mathcal{O}(\eta^2). \end{aligned}$$

Next, we substitute the expansion of $G_\alpha(x, \xi)$ into the expectation and variance terms with respect to $p(x \mid \xi)$, so that:

$$\begin{aligned} \mathbb{E}_{p(x \mid \xi)}[G_\alpha(x, \xi)] &= I(\xi) + \frac{\delta}{2}\, \mathbb{E}_{p(x \mid \xi)} \Big[\mathbb{V}_{p(\theta \mid x, \xi)} \big[L(\theta, x, \xi)\big]\Big] + \mathcal{O}(\delta^2), \\ \mathbb{V}_{p(x \mid \xi)}[G_\alpha(x, \xi)] &= \mathbb{V}_{p(x \mid \xi)} \Big[\mathbb{D}_{\mathrm{KL}}\big[p(\theta \mid x, \xi) \,\big\|\, p(\theta)\big]\Big] + \mathcal{O}(\delta), \end{aligned}$$

where we used $I(\xi) = \mathbb{E}_{p(x|\xi)}\left[\mathbb{D}_{\mathrm{KL}}\big[p(\theta \mid x, \xi) \;\big\|\; p(\theta)\big]\right]$. Substituting both back into $I^S_\alpha(\theta; x)(\xi)$:

$$I^S_\alpha(\theta; x)(\xi) = I(\xi) + \frac{\delta}{2}\mathbb{E}_{p(x|\xi)}\Big[\mathbb{V}_{p(\theta|x,\xi)}\left[L(\theta, x, \xi)\right]\Big] + \frac{\delta}{2}\mathbb{V}_{p(x|\xi)}\Big[\mathbb{E}_{p(\theta|x,\xi)}\left[L(\theta, x, \xi)\right]\Big] + \mathcal{O}(\delta^2),$$

where we used $\mathbb{D}_{\mathrm{KL}}\big[p(\theta \mid x, \xi) \;\big\|\; p(\theta)\big] = \mathbb{E}_{p(\theta|x,\xi)}\left[L(\theta, x, \xi)\right]$. Finally, by the law of total variance applied to L under the joint $p(\theta, x \mid \xi) = p(x \mid \xi)\, p(\theta \mid x, \xi)$:

$$\mathbb{V}_{p(\theta,x|\xi)}\big[L(\theta, x, \xi)\big] = \mathbb{E}_{p(x|\xi)}\Big[\mathbb{V}_{p(\theta|x,\xi)}\big[L(\theta, x, \xi)\big]\Big] + \mathbb{V}_{p(x|\xi)}\Big[\mathbb{E}_{p(\theta|x,\xi)}\left[L(\theta, x, \xi)\right]\Big].$$

Substituting and restoring $\delta = \alpha - 1$ yields the expansion:

$$I^S_\alpha(\theta; x)(\xi) = I(\xi) + \frac{\alpha - 1}{2}\,\mathbb{V}_{p(\theta,x|\xi)}\left[\log \frac{p(\theta \mid x, \xi)}{p(\theta)}\right] + \mathcal{O}\left((\alpha - 1)^2\right).$$

□

B.4 Proof of Proposition 2

For the proof that follows, we require that the nominal model carries finite information content.

Assumption 1 (Finite expected surprise)**.** *The integral*

$$\int p(\theta) \max\left\{1, p(x \mid \theta, \xi)\right\} \left|\log p(x \mid \theta, \xi)\right| \mathrm{d}\theta < \infty$$

is well-defined and finite for all (ξ, x)*. Additionally, the geometric average log-likelihood*

$$H_0(\xi) \coloneqq \int \exp\left\{\int p(\theta) \log p(x \mid \theta, \xi)\, \mathrm{d}\theta\right\} \mathrm{d}x \tag{10}$$

is uniformly bounded away from 0*: there exists a constant* $C_0 > 0$*, such that for all* $\xi \in \Xi$*, it holds:* $1 \geq H_0(\xi) > C_0$.

The behavior of the robust expected information gain is driven by the multiplier $\alpha/(\alpha - 1)$, so it is sufficient to verify that the remaining integral term remains bounded. For this purpose, we identify $H_0(\xi)$, defined in (10), as a uniform lower bound. This bound permits a direct application of Assumption 1, from which the desired convergence result follows.

Let us start by defining the pointwise quantity:

$$I(x, \xi; \alpha) \coloneqq \log\left\{\left[\int p(\theta)\, p(x \mid \theta, \xi)^\alpha \,\mathrm{d}\theta\right]^{1/\alpha}\right\}.$$

Note that $I(x, \xi; \cdot)$ is non-decreasing in α. We determine its limit as $\alpha \to 0$ by writing it as a ratio $f(\alpha)/\alpha$:

$$I(x, \xi; \alpha) = \frac{1}{\alpha} \log \int p(\theta)\, p(x \mid \theta, \xi)^\alpha \,\mathrm{d}\theta.$$

This presents an indeterminate limit $0/0$. Applying L'Hôpital's rule, the limit corresponds to $\lim_{\alpha \to 0^+} f'(\alpha)$, provided it exists. The derivative is:

$$f'(\alpha) = \frac{\dfrac{\partial}{\partial \alpha} \displaystyle\int p(\theta)\, p(x \mid \theta, \xi)^\alpha \,\mathrm{d}\theta}{\displaystyle\int p(\theta)\, p(x \mid \theta, \xi)^\alpha \,\mathrm{d}\theta}.$$

We analyze the denominator and numerator of $f'(\alpha)$ separately. For the denominator, Lebesgue's dominated convergence theorem, applied with the bound $p(x \mid \theta, \xi)^\alpha \leq \max\{1, p(x \mid \theta, \xi)\}$, ensures convergence to 1. For the numerator, the same theorem, combined with Lebesgue's differentiation theorem applied to the derivative of the integrand:

$$\left|p(x \mid \theta, \xi)^\alpha \log p(x \mid \theta, \xi)\right| \leq \max\left\{1, p(x \mid \theta, \xi)\right\} \left|\log p(x \mid \theta, \xi)\right|,$$

which is finite by Assumption 1, ensures that the numerator converges to

$$\int p(\theta) \log p(x \mid \theta, \xi)\, \mathrm{d}\theta.$$

Combining these results yields the pointwise bounds for all $\alpha \in (0,1)$, $x \in \mathcal{X}$, and $\xi \in \Xi$:

$$\int p(\theta) \log p(x \mid \theta, \xi)\, \mathrm{d}\theta \le I(x, \xi; \alpha) \le I(x, \xi; 1) = \log \int p(\theta)\, p(x \mid \theta, \xi)\, \mathrm{d}\theta.$$

Exponentiating and integrating with respect to x provides

$$H_0(\xi) \le \int \exp\big\{ I(x, \xi; \alpha) \big\}\, \mathrm{d}x \le \iint p(\theta)\, p(x \mid \theta, \xi)\, \mathrm{d}\theta\, \mathrm{d}x = 1.$$

By Assumption 1, we have $H_0(\xi) \ge C_0 > 0$. Consequently, the log-integral term is strictly bounded:

$$-\infty < \log C_0 \le \log \int \left[\int p(\theta)\, p(x \mid \theta, \xi)^{\alpha}\, \mathrm{d}\theta \right]^{1/\alpha} \mathrm{d}x \le 0.$$

This bound holds uniformly in ξ and α. Since the multiplier $\alpha/(\alpha - 1)$ converges to zero, the robust expected information gain vanishes as $\alpha \to 0$, completing the proof. □

B.5 Proof of Proposition 4

This assumption, used in Proposition 4 and Lemmata 3 and 4, controls the design objective behavior.

Assumption 2 (Regularity and boundedness)**.** *We assume the parameter $\alpha \in (0,1)$ and that the design set Ξ has a finite diameter D_Ξ. The ratios $w(x, \theta, \cdot)^\alpha$ are L_w-Lipschitz, bounded from below by $c_w > 0$ and from above by C_w, and have bounded variance $\sigma_w^2 \coloneqq \sup_{\xi, x} \mathbb{V}_\theta[w^\alpha]$. On the interval $[c_w, C_w]$, the function h is L_h-Lipschitz, bounded above by C_h, and has bounded variance $\sigma_h^2 \coloneqq \sup_\xi \mathbb{V}_{x|\xi}[(\mathbb{E}_\theta[w^\alpha])^{1/\alpha}]$. Consequently, the functions g and $\tilde{g}$ take values in $[\tau, C_h]$ with $\tau \coloneqq c_w^{1/\alpha} > 0$, on which the function f is L_f-Lipschitz.*

Lemma 3 (Bias bound for $\tilde{g}(\xi)$, Hu et al., 2020)**.** *Under Assumption 2, for any design $\xi \in \Xi$, the bias of the estimator $\tilde{g}(\xi)$ is bounded by:*

$$\big| \mathbb{E}\big[\tilde{g}(\xi)\big] - g(\xi) \big| \le \frac{L_h \sigma_w}{\sqrt{M}}.$$

Lemma 4 (Variance bound for $\tilde{g}(\xi)$, Hu et al., 2020)**.** *Under Assumption 2, for any design $\xi \in \Xi$, the variance of the estimator $\tilde{g}(\xi)$ is bounded by:*

$$\mathbb{V}\big[\tilde{g}(\xi)\big] \le \frac{\sigma_h^2}{N} + \frac{4 C_h L_h \sigma_w}{N\sqrt{M}}.$$

Recall that $\tilde{I}_\alpha^S(\xi) = f(\tilde{g}(\xi))$, where the functions f and g are defined in Section 4. We decompose the absolute error by adding and subtracting the term $f(\mathbb{E}[\tilde{g}])$ and then applying the triangle inequality:

$$\big| \mathbb{E}[f(\tilde{g})] - f(g) \big| \le \big| \mathbb{E}[f(\tilde{g})] - f(\mathbb{E}[\tilde{g}]) \big| + \big| f(\mathbb{E}[\tilde{g}]) - f(g) \big|.$$

We bound the first term by applying Jensen's inequality and invoking the Lipschitz continuity of f:

$$\begin{aligned} \big| \mathbb{E}[f(\tilde{g})] - f(\mathbb{E}[\tilde{g}]) \big| &= \big| \mathbb{E}[f(\tilde{g}) - f(\mathbb{E}[\tilde{g}])] \big| \\ &\le \mathbb{E}[|f(\tilde{g}) - f(\mathbb{E}[\tilde{g}])|] \le L_f\, \mathbb{E}[|\tilde{g} - \mathbb{E}[\tilde{g}]|], \end{aligned}$$

then we apply the Cauchy–Schwarz inequality, $\mathbb{E}[|X|] \le \sqrt{\mathbb{E}[X^2]}$, and obtain:

$$L_f\, \mathbb{E}\left[|\tilde{g} - \mathbb{E}[\tilde{g}]| \right] \le L_f \sqrt{\mathbb{V}[\tilde{g}]},$$

where the variance $\mathbb{V}[\tilde{g}]$ is controlled by Lemma 4. The second term can be bounded directly using Lipschitz continuity:

$$\big| f(\mathbb{E}[\tilde{g}]) - f(g) \big| \le L_f \big| \mathbb{E}[\tilde{g}] - g \big|,$$

where $\big| \mathbb{E}[\tilde{g}] - g \big|$ is bounded by Lemma 3. Finally, combining these bounds yields the stated result. □

B.6 Proof of Proposition 5

We briefly state a two-sided version of Hoeffding's inequality for bounded random variables that aids in the proof.

Lemma 5 (Two-sided Hoeffding inequality for bounded random variables, Vershynin, 2018)**.** *Let $X_1, \ldots, X_N$ be independent random variables such that $X_i \in [a_i, b_i]$ for every $i \in \{1, \ldots, N\}$. Then, for any $t > 0$, we have*

$$\mathbb{P}\Big(\Big| \sum_{i=1}^N \big(X_i - \mathbb{E}[X_i] \big) \Big| \ge t \Big) \le 2 \exp\left\{ - \frac{2t^2}{\sum_{i=1}^N (b_i - a_i)^2} \right\}.$$

Recall that $\tilde{I}_\alpha^S(\xi) = f(\tilde{g}(\xi))$, where the functions f and g are defined in Section 4. By invoking the Lipschitz continuity of f and applying the triangle inequality, we decompose the total error into a *stochastic* and *deterministic* component:

$$\begin{aligned}\left|\tilde{I}_\alpha^S - I_\alpha^S\right| = \left|f(\tilde{g}) - f(g)\right| &\leq L_f\left|\tilde{g} - g\right| \\ &\leq \underbrace{L_f\left|\tilde{g} - \mathbb{E}[\tilde{g}]\right|}_{\text{stochastic}} + \underbrace{L_f\left|\mathbb{E}[\tilde{g}] - g\right|}_{\text{deterministic}}.\end{aligned}$$

According to Lemma 3, the deterministic term $\left|\mathbb{E}[\tilde{g}] - g\right|$ is upper bounded by $L_h\sigma_w/\sqrt{M}$. To ensure the total error stays within tolerance t, we constrain this error term to consume at most half of the total budget $t/2$ by choosing M such that:

$$\frac{L_h\sigma_w}{\sqrt{M}} \leq \frac{t}{2L_f} \quad \implies M \geq \frac{4L_f^2 L_h^2 \sigma_w^2}{t^2}.$$

Consequently, for the total error to exceed t, the stochastic term $\left|\tilde{g} - \mathbb{E}[\tilde{g}]\right|$ must account for the remaining difference:

$$L_f\left|\tilde{g} - \mathbb{E}[\tilde{g}]\right| \geq t - \frac{t}{2} = \frac{t}{2} \quad \implies \quad \left|\tilde{g} - \mathbb{E}[\tilde{g}]\right| \geq \frac{t}{2L_f}.$$

Next, we bound the probability of this deviation. Recall that $\tilde{g}$ is the empirical average of N independent random variables $\tilde{h}^{(i)}$, defined as:

$$\tilde{h}^{(i)} \coloneqq h\Big(\frac{1}{M}\sum_{j=1}^{M} w(x^{(i)}, \theta^{(i,j)}, \xi)^\alpha\Big).$$

By Assumption 2, the function h is bounded by C_h, implying that each random variable $\tilde{h}^{(i)} \in [0, C_h]$. We can therefore apply Lemma 5 to the sample average $\tilde{g} = \frac{1}{N}\sum_{i=1}^{N}\tilde{h}^{(i)}$:

$$\mathbb{P}\Big(\left|\tilde{g} - \mathbb{E}[\tilde{g}]\right| \geq \frac{t}{2L_f}\Big) \leq 2\exp\left\{-\frac{Nt^2}{2L_f^2 C_h^2}\right\}.$$

Combining the bounds on the stochastic and deterministic components, we obtain the final concentration inequality:

$$\mathbb{P}\left(\left|\tilde{I}_\alpha^S - I_\alpha^S\right| \geq t\right) \leq 2\exp\left\{-\frac{Nt^2}{2L_f^2 C_h^2}\right\}.$$

□

B.7 Proof of Proposition 6

This proof follows standard PAC-Bayes proof techniques presented by Alquier [2024]. We begin by stating three useful lemmata that will facilitate the derivation.

Lemma 6 (Hoeffding's lemma, Vershynin, 2018)**.** *Let X be a random variable such that $X \in [a, b]$ almost surely. Then, for any $s \in \mathbb{R}$, it holds that:*

$$\mathbb{E}\left[\exp\{s\,(X - \mathbb{E}[X])\}\right] \leq \exp\left\{\frac{s^2(b-a)^2}{8}\right\}.$$

Lemma 7 (Markov inequality, Vershynin, 2018)**.** *For any non-negative random variable X and $a > 0$, we have:*

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a} \quad \iff \quad \mathbb{P}(X \geq k\,\mathbb{E}[X]) \leq \frac{1}{k}, \text{ where } k = \frac{a}{\mathbb{E}[X]}.$$

Lemma 8 (Donsker–Varadhan variational formula, Donsker and Varadhan, 1975)**.** *For any measurable, bounded function $f : \Xi \mapsto \mathbb{R}$, and any probability measure $\pi_0 \in \Pi$, we have:*

$$\log \mathbb{E}_{\pi_0}\left[e^{f(\xi)}\right] = \sup_{\pi \in \Pi}\left\{\mathbb{E}_\pi\left[f(\xi)\right] - \mathbb{D}_{\mathrm{KL}}[\pi \parallel \pi_0]\right\}.$$

To prove the proposition, we first analyze the concentration properties of the estimation error $\tilde{I}_\alpha^S(\xi) - I_\alpha^S(\xi)$ for a fixed design $\xi \in \Xi$. Let $\bar{g}(\xi) \coloneqq \mathbb{E}[\tilde{g}(\xi)]$ denote the mean of the empirical

estimator, where the expectation is evaluated under the sample distribution of the naive nested Monte Carlo estimator. By the Lipschitz continuity of f under Assumption 2 and the bias bound of Lemma 3, the estimation error decomposes into two terms:

$$\tilde{I}_\alpha^S(\xi) - I_\alpha^S(\xi) = \big[f(\tilde{g}) - f(\bar{g})\big] + \big[f(\bar{g}) - f(g)\big] \leq L_f \left|\tilde{g} - \bar{g}\right| + \frac{L_f L_h \sigma_w}{\sqrt{M}}.$$

Recall from the proof of Proposition 5 that $\tilde{g}$ is the empirical average of N independent random variables $\tilde{h}^{(i)} \in [0, C_h]$. Using Lemma 6 on each $\tilde{h}^{(i)}$, for any $s \in \mathbb{R}$, we have:

$$\mathbb{E}\Big[\exp\big\{s\,(\tilde{g} - \bar{g})\big\}\Big] = \prod_{i=1}^{N} \mathbb{E}\Big[\exp\Big\{\frac{s}{N}\big(\tilde{h}^{(i)} - \mathbb{E}[\tilde{h}^{(i)}]\big)\Big\}\Big] \leq \exp\left\{\frac{s^2 C_h^2}{8N}\right\}.$$

Applying this for s and $-s$, and using the inequality $e^{s|z|} \leq e^{sz} + e^{-sz}$, we obtain for any $s > 0$:

$$\mathbb{E}\Big[\exp\big\{s\,|\tilde{g} - \bar{g}|\big\}\Big] \leq 2\exp\left\{\frac{s^2 C_h^2}{8N}\right\}.$$

Combining both results with $s = \lambda L_f$, for any $\lambda > 0$, the moment generating function of the estimation error is bounded as:

$$\mathbb{E}\Big[\exp\Big\{\lambda\Big(\tilde{I}_\alpha^S(\xi) - I_\alpha^S(\xi)\Big)\Big\}\Big] \leq 2\exp\left\{\frac{\lambda L_f L_h \sigma_w}{\sqrt{M}} + \frac{\lambda^2 L_f^2 C_h^2}{8N}\right\}.$$

We integrate this bound under the policy π_0 and swap the expectations yielding:

$$\mathbb{E}\Big[\mathbb{E}_{\pi_0}\Big[\exp\Big\{\lambda\Big(\tilde{I}_\alpha^S(\xi) - I_\alpha^S(\xi)\Big)\Big\}\Big]\Big] \leq 2\exp\left\{\frac{\lambda L_f L_h \sigma_w}{\sqrt{M}} + \frac{\lambda^2 L_f^2 C_h^2}{8N}\right\}.$$

Next, we apply Lemma 8 to the inner expectation and obtain the following:

$$\mathbb{E}\Big[\exp\Big\{\sup_{\pi\in\Pi}\Big\{\lambda\,\mathbb{E}_\pi\Big[\tilde{I}_\alpha^S(\xi) - I_\alpha^S(\xi)\Big] - \mathbb{D}_{\mathrm{KL}}[\pi \parallel \pi_0]\Big\}\Big\}\Big] \leq 2\exp\left\{\frac{\lambda L_f L_h \sigma_w}{\sqrt{M}} + \frac{\lambda^2 L_f^2 C_h^2}{8N}\right\}.$$

This gives us a bound on the expectation of the form: $\mathbb{E}[X] \leq b$, where X is the exponential term inside the outer expectation. To convert this expectation bound back to a probability bound, we apply Lemma 7 on X and assume $k = 1/\delta$ for $\delta \in (0, 1)$:

$$\mathbb{P}\Big(\exp\Big\{\sup_{\pi\in\Pi}\Big\{\lambda\,\mathbb{E}_\pi\big[\tilde{I}_\alpha^S(\xi) - I_\alpha^S(\xi)\big] - \mathbb{D}_{\mathrm{KL}}[\pi \parallel \pi_0]\Big\}\Big\} \geq \frac{2}{\delta}\exp\left\{\frac{\lambda L_f L_h \sigma_w}{\sqrt{M}} + \frac{\lambda^2 L_f^2 C_h^2}{8N}\right\}\Big) \leq \delta.$$

Taking the logarithm of both sides and rearranging the terms leads to:

$$\mathbb{P}\Big(\exists\pi \in \Pi,\, \mathbb{E}_\pi\big[\tilde{I}_\alpha^S(\xi)\big] - \frac{\lambda L_f^2 C_h^2}{8N} - \frac{L_f L_h \sigma_w}{\sqrt{M}} - \frac{\mathbb{D}_{\mathrm{KL}}[\pi \parallel \pi_0] + \log(2/\delta)}{\lambda} \geq \mathbb{E}_\pi\big[I_\alpha^S(\xi)\big]\Big) \leq \delta.$$

Finally, the condition $M \geq 2NL_h^2\sigma_w^2/(C_h^2\log(2/\delta))$ ensures that the bias term is dominated by the concentration terms. It implies $L_h\sigma_w/\sqrt{M} \leq C_h\sqrt{\log(2/\delta)/(2N)}$, so that by the inequality $2\sqrt{ab} \leq \lambda a + b/\lambda$, valid for any $a, b \geq 0$ and $\lambda > 0$, we have:

$$\frac{L_f L_h \sigma_w}{\sqrt{M}} \leq L_f C_h\sqrt{\frac{\log(2/\delta)}{2N}} \leq \frac{\lambda L_f^2 C_h^2}{8N} + \frac{\log(2/\delta)}{\lambda}.$$

Substituting this bound and taking the complement of the previous result, we retrieve a lower bound on $\mathbb{E}_\pi\left[I_\alpha^S(\xi)\right]$:

$$\mathbb{P}\Big(\forall\pi \in \Pi,\, \mathbb{E}_\pi\big[\tilde{I}_\alpha^S(\xi)\big] - \frac{\lambda L_f^2 C_h^2}{4N} - \frac{\mathbb{D}_{\mathrm{KL}}[\pi \parallel \pi_0] + 2\log(2/\delta)}{\lambda} \leq \mathbb{E}_\pi\big[I_\alpha^S(\xi)\big]\Big) \geq 1 - \delta,$$

that holds with probability at least $1 - \delta$. □

C Implementation of the Nested Monte Carlo Estimator

To ensure numerical stability and prevent arithmetic underflow, we perform all computations strictly in the logarithmic domain. We define the log-likelihood function as $\ell(\theta, x) \coloneqq \log p(x \mid \theta, \xi)$ and the log-sum-exp operator as $\mathrm{LSE}(y_{1:K}) \coloneqq \log\sum_{k=1}^{K}\exp(y_k)$. The estimation procedure proceeds in four steps:

- We first generate an outer batch of N independent (θ, x) pairs from the joint generative model. For each outer sample with index $i \in \{1, \dots, N\}$, we independently draw a set of M inner parameters from the prior:

$$\{(\theta^{(i)}, x^{(i)})\}_{i=1}^{N} \sim p(\theta)\, p(x \mid \theta, \xi), \qquad \{\theta^{(i,j)}\}_{j=1}^{M} \sim p(\theta) \quad \text{for each } i \in \{1, \dots, N\}.$$

- As per Corollary 2, we now need to compute $w(x^{(i)}, \theta^{(i,j)}, \xi) = p(x^{(i)} \mid \theta^{(i,j)}, \xi)/p(x^{(i)} \mid \xi) =: w^{(i,j)}$. When the marginal likelihood $p(x \mid \xi)$ is available in closed form, as in the conjugate models of Appendices F.1 and F.2, we evaluate it directly:

$$\log w^{(i,j)} = \ell(\theta^{(i,j)}, x^{(i)}) - \log p(x^{(i)} \mid \xi).$$

  When $p(x \mid \xi)$ is intractable, we instead use the contrastive estimator $\tilde{w}$ of Definition 3.

- We next estimate the inner expectation term $\varphi(x, \xi) = \mathbb{E}_{p(\theta)}[w(x, \theta, \xi)^{\alpha}]$ by averaging the powered weights. In the log-domain, this corresponds to a second application of the LSE operator over the M inner samples:

$$\log \varphi(x^{(i)}, \xi) = \mathrm{LSE}\left(\left\{\alpha \log w^{(i,j)}\right\}_{j=1}^{M}\right) - \log M.$$

- Finally, we compute the robust expected information gain by aggregating the outer samples. The estimator $\tilde{I}_{\alpha}^{S}(\xi)$ approximates the logarithm of the expected exponential gain:

$$\tilde{I}_{\alpha}^{S}(\xi) = \frac{\alpha}{\alpha - 1}\left[\mathrm{LSE}\left(\left\{\alpha^{-1} \log \varphi(x^{(i)}, \xi)\right\}_{i=1}^{N}\right) - \log N\right].$$

This procedure yields a consistent, albeit biased, estimate of the robust objective.

D Theoretical Comparison of PAC-Bayesian and Deterministic Policies

The theoretical advantages of PAC-Bayesian bounds are comprehensively summarized in Alquier [2024]. In the context of bandit problems, which are highly relevant to experimental design, stochastic policies derived from the PAC-Bayesian formulation facilitate efficient exploration while streamlining the theoretical analysis [Flynn et al., 2023]. This section aims to compare the concentration properties of deterministic and stochastic policies, illustrating that PAC-Bayesian bounds can yield a sharper upper estimate of the minimum of the information gain under certain conditions. These results further reinforce the preference for stochastic policies, as they do not necessarily compromise the concentration rate in spite of the induced randomness, while leveraging the theoretical versatility of the PAC-Bayesian framework.

Consider the minimization of an objective function $J(\xi)$ over the design space Ξ, where the analysis applies symmetrically to maximization. To streamline the illustration, we adopt a simplified setup where the objective function $J(\xi)$ is defined by $J(\xi) = \mathbb{E}_{z \sim p}[h(\xi, z)]$, where $h(\xi, \cdot)$ is an integrand with a latent variable z defined on a compact support $\mathcal{Z} \subset \mathbb{R}^k$ and p is the distribution of the latent variable z. We approximate the objective function $J(\xi)$ using the empirical Monte Carlo estimator $\tilde{J}(\xi)$ defined by N i.i.d. realizations $z_1, \dots, z_N$ drawn from p:

$$J(\xi) = \mathbb{E}_{z \sim p}[h(\xi, z)] \quad \approx \quad \tilde{J}(\xi) = \frac{1}{N}\sum_{i=1}^{N} h(\xi, z_i).$$

While we seek to minimize $J(\xi)$, the exact objective $J(\xi)$ is inaccessible in practice. Consequently, the deterministic policy minimizes the Monte Carlo estimator $\tilde{J}(\xi)$ as a proxy. We define

$$\xi_{\min} \in \arg\min_{\xi} J(\xi) \quad \text{and} \quad \tilde{\xi}_{\min} \in \arg\min_{\xi} \tilde{J}(\xi).$$

One justification for employing the deterministic policy is that the minimum of the Monte Carlo estimator $\tilde{J}(\xi)$ provides an upper estimate of the global minimum $J(\xi_{\min})$.

To derive the upper estimate, we introduce a standard measure from statistical learning theory, called the Rademacher complexity. Intuitively, the Rademacher complexity measures the expressive power of a function family relative to samples of size N. Let $\mathcal{R}_N$ denote the Rademacher complexity of the integrand family $\{h(\cdot, z) \mid z \in \mathcal{Z}\}$ computed over N random samples $\{z_i\}_{i=1}^{N}$ following p. For the

purposes of this illustration, the formal definition of Rademacher complexity is omitted for brevity. We refer the reader to Wainwright [2019] for a rigorous treatment of Rademacher complexity and its upper bounds for common function classes. The next assumption is used throughout this section.

Assumption 3. *The integrand $h(\xi, z)$ is non-negative, continuous, and uniformly bounded by a constant b. The objective function $J(\xi)$ is uniquely minimized over Ξ.*

The following upper estimate of the global minimum $J(\xi_{\min})$ is a standard result derived from empirical process theory, particularly in the study of empirical risk minimization.

Proposition 7. *It holds for any $N > 0$ and $\delta > 0$ that*

$$\mathbb{P}\left(J(\xi_{min}) \leq \min_{\xi} \tilde{J}(\xi) + 2\mathcal{R}_N + b\sqrt{\frac{2\log(1/\delta)}{N}} \right) \geq 1 - \delta,$$

where the minimum is attained at $\xi = \tilde{\xi}_{min}$.

Proof. We have $J(\xi_{\min}) \leq J(\xi)$ for any point ξ, since $\xi_{\min}$ is the minimum. This, in turn, yields that

$$J(\xi_{\min}) - \tilde{J}(\tilde{\xi}_{\min}) \leq J(\tilde{\xi}_{\min}) - \tilde{J}(\tilde{\xi}_{\min}) \leq \sup_{\xi} |J(\xi) - \tilde{J}(\xi)|.$$

The following concentration inequality is a direct result of Theorem 4.10 of Wainwright [2019]:

$$\mathbb{P}\left(\sup_{\xi} |J(\xi) - \tilde{J}(\xi)| \leq 2\mathcal{R}_N + b\sqrt{\frac{2\log(1/\delta)}{N}} \right) \geq 1 - \delta.$$

Rearranging the term completes the proof. □

By definition, the empirical minimizer $\tilde{\xi}_{\min}$ yields the sharpest possible bound of this form, as any deviation from the empirical minimum necessarily increases the value of the upper estimate.

In contrast to the above concentration inequality, the PAC-Bayesian framework provides a flexible upper estimate that bypasses the need for global complexity measures such as the Rademacher complexity. Given a prior π_0 over the design space Ξ, we define a stochastic (Gibbs) policy $\pi^\star$ as

$$\pi^\star(\xi) \propto \exp\left\{-\lambda \tilde{J}(\xi)\right\} \pi_0(\xi) = \arg\min_{\rho} \left\{ \mathbb{E}_\rho[\tilde{J}(\xi)] + \frac{\mathbb{D}_{\mathrm{KL}}[\rho \parallel \pi_0]}{\lambda} \right\},$$

where $\lambda > 0$ is the precision hyperparameter that controls the policy dispersion. A well-established bound by Catoni [2007] provides the following upper estimate under the stochastic policy $\pi^\star$.

Proposition 8. *It holds for any $N > 0$ and $\delta > 0$ that*

$$\mathbb{P}\left(J(\xi_{min}) \leq \min_{\rho} \left\{ \mathbb{E}_\rho[\tilde{J}(\xi)] + \frac{\mathbb{D}_{\mathrm{KL}}[\rho \parallel \pi_0]}{\lambda} \right\} + \frac{\lambda b^2}{8N} + \frac{\log(1/\delta)}{\lambda} \right) \geq 1 - \delta,$$

where the minimum is attained at $\rho = \pi^\star$.

Proof. We have $J(\xi_{\min}) \leq \mathbb{E}_\rho[J(\xi)]$ for an arbitrary distribution ρ, since $\xi_{\min}$ is the minimum and J is non-negative. Applying Catoni's bound [Catoni, 2007] to $\mathbb{E}_\rho[J(\xi)]$ completes the proof. □

Employing the stochastic policy offers two primary benefits: (i) it permits a more tractable and flexible theoretical treatment by circumventing dependency on global complexity measures, and (ii) it promotes exploration across the design space Ξ, as discussed in Flynn et al. [2023]. Furthermore, it may yield a sharper upper estimate of the global minimum $J(\xi_{\min})$ than the standard bound. Finally, the following proposition demonstrates that the PAC-Bayesian bound can be strictly sharper than the other, provided that the level δ and the precision parameter λ are appropriately chosen.

Proposition 9. *Choose a level δ such that $\delta \leq e^{-8N}$ for a given value of the sample number N. Choose a precision parameter λ such that $\lambda \in ((3-\sqrt{5})\sqrt{2(N/b^2)\log(1/\delta)}, (3+\sqrt{5})\sqrt{2(N/b^2)\log(1/\delta)})$. Then, it holds that*

$$\min_{\rho} \left\{ \mathbb{E}_\rho[\tilde{J}(\xi)] + \frac{\mathbb{D}_{\mathrm{KL}}[\rho \parallel \pi_0]}{\lambda} \right\} + \frac{\lambda b^2}{8N} + \frac{\log(1/\delta)}{\lambda} < \min_{\xi} \tilde{J}(\xi) + 2\mathcal{R}_N + b\sqrt{\frac{2\log(1/\delta)}{N}}.$$

Proof. By the Donsker–Varadhan variational formula from Lemma 8, we have

$$\min_{\rho}\left\{\mathbb{E}_{\rho}[\tilde{J}(\xi)]+\frac{\mathbb{D}_{\mathrm{KL}}[\rho\parallel\pi_0]}{\lambda}\right\}=-\frac{1}{\lambda}\log\left(\mathbb{E}_{\pi_0}[\exp\{-\lambda\tilde{J}(\xi)\}]\right).$$

It follows from the upper bound of δ that $b\le b(1/4)\sqrt{2\log(1/\delta)/N}$. Since $\tilde{J}$ is non-negative and bounded by b, we have $\tilde{J}(\xi)\le b\le\tilde{J}(\tilde{\xi}_{\min})+2\mathcal{R}_N+(1/4)b\sqrt{2\log(1/\delta)/N}$. Substituting this into the Donsker–Varadhan formula gives that

$$\min_{\rho}\left\{\mathbb{E}_{\rho}[\tilde{J}(\xi)]+\frac{\mathbb{D}_{\mathrm{KL}}[\rho\parallel\pi_0]}{\lambda}\right\}\le\tilde{J}(\tilde{\xi}_{\min})+2\mathcal{R}_N+\frac{b}{4}\sqrt{\frac{2\log(1/\delta)}{N}}.$$

Thus, to establish the main result, it suffices to show that

$$\frac{\lambda b^2}{8N}+\frac{\log(1/\delta)}{\lambda}\le\frac{3b}{4}\sqrt{\frac{2\log(1/\delta)}{N}}\quad\Longleftrightarrow\quad\frac{b^2}{8N}\lambda^2-\frac{3b}{4}\sqrt{\frac{2\log(1/\delta)}{N}}\lambda+\log\frac{1}{\delta}\le 0.$$

Solving the quadratic inequality on the right hand side completes the proof. □

This comparison is not intended to suggest that the PAC-Bayesian framework generally yields a sharper bound than the classical counterparts; their performance remains contingent on the choice of prior and the concentration of the empirical objective. Nevertheless, it illustrates that the PAC-Bayesian framework can provide the latitude to enhance concentration properties, while offering flexibility in the theoretical analysis. In this illustration, for some confidence level δ, there exists a regime of the precision parameter λ where the PAC-Bayesian bound is strictly sharper than the other bound. Such theoretical insights may further facilitate a principled selection of the precision parameter λ in practice.

E Choosing the PAC-Bayes Precision λ

The parameter λ controls the inverse temperature of the PAC-Bayes policy and the tightness of the finite-sample bound. For a fixed posterior policy π, the penalty term in Proposition 6 has the form:

$$\frac{L_f^2\,C_h^2}{4N}\lambda+\frac{\mathbb{D}_{\mathrm{KL}}[\pi\parallel\pi_0]+2\log(2/\delta)}{\lambda},$$

and is minimized by the oracle choice

$$\lambda_{\text{oracle}}(\pi)=\sqrt{\frac{4N\left(\mathbb{D}_{\mathrm{KL}}[\pi\parallel\pi_0]+2\log(2/\delta)\right)}{L_f^2C_h^2}}.$$

This expression is useful for interpreting the estimation–complexity trade-off of the bound. The optimal precision $\lambda_{\text{oracle}}(\pi)$ grows with the Monte Carlo budget N. Intuitively, as N grows, the Monte Carlo error in $\tilde{I}_\alpha^S$ decreases, so the empirical robust EIG can be trusted more and the Gibbs policy concentrates more strongly around designs with high empirical robust EIG. At the same time, larger N reduces the cost of the estimation penalty $\lambda L_f^2C_h^2/(4N)$, allowing a larger inverse temperature before the finite-sample correction becomes too large. Since the KL term is divided by λ, this also means that, for a fixed complexity $\mathbb{D}_{\mathrm{KL}}[\pi\parallel\pi_0]$, larger Monte Carlo budgets permit larger deviations from the prior policy π_0 while maintaining a valid certificate.

Conversely, Proposition 5 shows that L_f and C_h control the concentration of the Monte Carlo estimator. Larger values of either quantity negatively impact the estimator's reliability. To compensate, the optimal λ shrinks, forcing the policy to stay closer to the prior to prevent overfitting to Monte Carlo noise. Overall, the choice of λ should increase when the empirical estimator is reliable, and decrease when the estimator is more variable.

Because λ_{oracle} depends on the learned policy π, itself learned under Monte Carlo samples, it cannot be computed directly. Thankfully, the PAC-Bayes bound in Proposition 6 is valid for any value of λ fixed independently of those samples. To tune λ, we can instead follow the standard PAC-Bayesian treatment described by Alquier [2024].

To this end, consider a finite grid $\Lambda\subset(0,\infty)$, chosen before observing the Monte Carlo samples. Using the same argument as Proposition 6, applied separately to each fixed $\lambda\in\Lambda$ with confidence

Table 5: Sensitivity of PAC-Bayes policies to the choice of λ on the linear regression problem. We sweep over λ and N and report the mean relative regret.

N	$\lambda = 10^2$	$\lambda = 10^3$	$\lambda = 10^4$	$\lambda = 10^5$	$\lambda = 10^6$	$\lambda = 10^7$	$\lambda = 10^8$
64	0.062	0.020	0.008	**0.011**	**0.026**	**0.023**	0.019
128	0.081	0.037	0.037	**0.024**	**0.006**	**0.013**	0.023
256	0.094	0.019	0.020	**0.021**	**0.019**	**0.022**	0.018

level $\delta/|\Lambda|$ and taking a union bound over the grid, we can prove that, with probability at least $1-\delta$, for all $\pi \in \Pi$ and all $\lambda \in \Lambda$ we have:

$$\mathbb{E}_\pi\big[I^S_\alpha(\xi)\big] \geq \mathbb{E}_\pi\big[\tilde{I}^S_\alpha(\xi)\big] - \frac{\lambda L_f^2 C_h^2}{4N} - \frac{\mathbb{D}_{\mathrm{KL}}[\pi \parallel \pi_0] + 2\log(2|\Lambda|/\delta)}{\lambda}.$$

The choice of λ is now clear: it may be selected together with the policy by maximizing the grid-based lower bound over λ, π:

$$\pi^\star, \lambda^\star = \arg\max_{\pi\in\Pi, \lambda\in\Lambda} \mathbb{E}_\pi\big[\tilde{I}^S_\alpha(\xi)\big] - \frac{\lambda L_f^2 C_h^2}{4N} - \frac{\mathbb{D}_{\mathrm{KL}}[\pi \parallel \pi_0] + 2\log(2|\Lambda|/\delta)}{\lambda}.$$

This is done by simple exhaustion: for each $\lambda \in \Lambda$, one computes the Gibbs policy $\pi_\lambda(\xi) \propto \pi_0(\xi)\exp\big\{\lambda\, \tilde{I}^S_\alpha(\xi)\big\}$ and selects the value of λ whose lower bound is largest.

In practice, evaluating the PAC-Bayes bound exactly may be infeasible when the constants L_f and C_h are unknown. To assess how sensitive the policy optimization problem in (8) is to the choice of λ, we can empirically study λ, using the regret of the learned policy π_λ relative to a tractable Sibson's α-MI benchmark in the linear regression setting.

To do so, we fix the inner sample size $M = 16$ and perform a joint sweep over $\lambda \in \{10^2, \ldots, 10^8\}$ and the outer sample size $N \in \{64, 128, 256\}$ on the high-dimensional linear regression problem. For each pair (N, λ), we learn a separate PAC-Bayes policy. Table 5 reports the mean relative regret over $4\,096$ design samples of the corresponding learned policy. While individual entries are noisy, partly due to the stochastic mirror-descent optimizer, the results exhibit a stable low-regret plateau for $\lambda \in [10^5, 10^7]$, and this pattern is consistent across all tested values of N. Although this sweep is not intended as a formal tuning procedure for λ, its results are consistent with the evaluations in Section 7, where policies learned with $\lambda = 10^6$ achieve substantially lower regret than a naive optimizer.

F Evaluation Details

All numerical evaluations reported in this paper require negligible compute and memory: every experiment runs on a standard CPU and completes within minutes.

F.1 Linear Regression

To illustrate the robust experimental design framework, we derive closed-form expressions for the worst-case distribution and robust expected information gain in the conjugate linear regression setting. This model admits tractable analytical results that provide intuition for the general case.

Consider the standard Bayesian linear regression model. Let $\theta = (\theta_1, \theta_2)^\top \in \mathbb{R}^2$ denote the unknown parameter vector, where $\theta_1 \in \mathbb{R}$ is the slope and $\theta_2 \in \mathbb{R}$ is the offset. We endow θ with a Gaussian prior $p(\theta) = \mathrm{N}(\theta; \mu_0, \Sigma_0)$, where $\mu_0 \in \mathbb{R}^2$ is the prior mean and $\Sigma_0 \in \mathbb{R}^{2\times 2}$ is the prior covariance matrix. A batch design $\xi_{1:N} = (\xi_1, \ldots, \xi_N) \in \Xi$ specifies N individual designs $\xi_i \in \mathbb{R}$. We construct the augmented design matrix $H(\xi_{1:N}) \in \mathbb{R}^{N\times 2}$ by appending a column of ones:

$$H(\xi_{1:N}) = \begin{bmatrix} \xi_1 & 1 \\ \vdots & \vdots \\ \xi_N & 1 \end{bmatrix}.$$

Conditioned on the parameter θ and design $\xi_{1:N}$, the N measurements are collected as a vector $x_{1:N} = (x_1, \ldots, x_N)^\top \in \mathbb{R}^N$, where each $x_i \in \mathbb{R}$ is a scalar observation. Under the assumption of

independence, the likelihood factorizes as:

$$p(x_{1:N} \mid \theta, \xi_{1:N}) = \prod_{i=1}^{N} p(x_i \mid \theta, \xi_i) = \prod_{i=1}^{N} \mathrm{N}(x_i; \xi_i\,\theta_1 + \theta_2, \sigma^2) = \mathrm{N}(x_{1:N}; H(\xi_{1:N})\,\theta, \sigma^2 I_N),$$

where $\sigma^2 > 0$ is the observation noise variance.

Marginal and Posterior. Under these conjugate assumptions, the marginal likelihood admits closed-form Gaussian expressions:

$$p(x_{1:N} \mid \xi_{1:N}) = \mathrm{N}(x_{1:N}; H(\xi_{1:N})\,\mu_0, H(\xi_{1:N})\,\Sigma_0\,H(\xi_{1:N})^\top + \sigma^2 I_N).$$

Additionally, the posterior $p(\theta \mid x_{1:N}, \xi_{1:N}) = \mathrm{N}(\theta; \mu_N, \Sigma_N)$ is given by the standard conjugate update formulas:

$$\Sigma_N = \left[\Sigma_0^{-1} + \frac{1}{\sigma^2} H(\xi_{1:N})^\top H(\xi_{1:N})\right]^{-1}, \quad \mu_N = \Sigma_N \left[\Sigma_0^{-1}\mu_0 + \frac{1}{\sigma^2} H(\xi_{1:N})^\top x_{1:N}\right].$$

Worst-Case Distribution. The worst-case joint distribution $q^\star(\theta, x_{1:N} \mid \xi_{1:N})$ is given by:

$$q^\star(\theta, x_{1:N} \mid \xi_{1:N}) \propto \Big[p(\theta)\,p_\alpha(x_{1:N} \mid \xi_{1:N})\Big]^{1-\alpha} \Big[p(\theta, x_{1:N} \mid \xi_{1:N})\Big]^{\alpha},$$

where the tilted marginal $p_\alpha(x_{1:N} \mid \xi_{1:N})$ is defined as:

$$p_\alpha(x_{1:N} \mid \xi_{1:N}) \propto \left[\mathbb{E}_{p(\theta)}\Big[p(x_{1:N} \mid \theta, \xi_{1:N})^\alpha\Big]\right]^{1/\alpha}.$$

For this conjugate problem, the α-powered likelihood is:

$$p(x_{1:N} \mid \theta, \xi_{1:N})^\alpha = (2\pi\sigma^2)^{-\alpha N/2} \exp\left\{-\frac{\alpha}{2\sigma^2}\|x_{1:N} - H(\xi_{1:N})\,\theta\|^2\right\},$$

where $\|x_{1:N} - H(\xi_{1:N})\,\theta\|^2 = \sum_{i=1}^{N}(x_i - \xi_i\theta_1 - \theta_2)^2$. To compute the expectation $\mathbb{E}_{p(\theta)}\big[p(x_{1:N} \mid \theta, \xi_{1:N})^\alpha\big]$, we expand the quadratic form and integrate over θ. Completing the square in θ yields:

$$\begin{aligned}\mathbb{E}_{p(\theta)}\Big[p(x_{1:N} \mid \theta, \xi_{1:N})^\alpha\Big] &= (2\pi\sigma^2)^{-\alpha N/2} \int \exp\left\{-\frac{\alpha}{2\sigma^2}\|x_{1:N} - H(\xi_{1:N})\,\theta\|^2\right\} p(\theta)\;\mathrm{d}\theta \\ &= (2\pi\sigma^2)^{-\alpha N/2} \frac{|\Psi_\alpha|^{1/2}}{|\Sigma_0|^{1/2}} \exp\left\{-\frac{1}{2}\frac{\alpha}{\sigma^2}\|x_{1:N}\|^2 - \frac{1}{2}\mu_0^\top\Sigma_0^{-1}\mu_0 + \frac{1}{2}m_\alpha^\top\Psi_\alpha^{-1}m_\alpha\right\},\end{aligned}$$

where we define the tilted covariance matrix and mean:

$$\Psi_\alpha \coloneqq \left[\Sigma_0^{-1} + \frac{\alpha}{\sigma^2} H(\xi_{1:N})^\top H(\xi_{1:N})\right]^{-1}, \quad m_\alpha \coloneqq \Psi_\alpha \left[\Sigma_0^{-1}\mu_0 + \frac{\alpha}{\sigma^2} H(\xi_{1:N})^\top x_{1:N}\right].$$

After simplification the tilted marginal takes the form:

$$p_\alpha(x_{1:N} \mid \xi_{1:N}) \propto \exp\left\{-\frac{1}{2}\big[x_{1:N} - H(\xi_{1:N})\,\mu_0\big]^\top (\alpha\Lambda_\alpha)^{-1}\big[x_{1:N} - H(\xi_{1:N})\,\mu_0\big]\right\},$$

where Λ_α is the α-scaled predictive covariance for the batch:

$$\Lambda_\alpha(\xi_{1:N}) \coloneqq H(\xi_{1:N})\,\Sigma_0\,H(\xi_{1:N})^\top + \frac{\sigma^2}{\alpha}\,I_N.$$

Therefore, the tilted marginal is Gaussian:

$$p_\alpha(x_{1:N} \mid \xi_{1:N}) = \mathrm{N}\big(x_{1:N}; H(\xi_{1:N})\,\mu_0, \alpha\Lambda_\alpha(\xi_{1:N})\big).$$

Overall, this results in a worst-case joint distribution $q^\star(\theta, x_{1:N} \mid \xi_{1:N})$ that is also Gaussian:

$$q^\star(\theta, x_{1:N} \mid \xi_{1:N}) = \mathrm{N}\left(\begin{bmatrix}\theta \\ x_{1:N}\end{bmatrix}; \mu_q^\star, \Sigma_q^\star\right),$$

with mean and covariance:

$$\mu_q^\star = \begin{bmatrix}\mu_0 \\ H(\xi_{1:N})\,\mu_0\end{bmatrix}, \quad \Sigma_q^\star = \begin{bmatrix}\alpha\Sigma_0 + (1-\alpha)\,\Sigma_\theta^\star & \alpha\Sigma_0\,H(\xi_{1:N})^\top \\ \alpha H(\xi_{1:N})\,\Sigma_0 & \alpha\Lambda_\alpha(\xi_{1:N})\end{bmatrix},$$

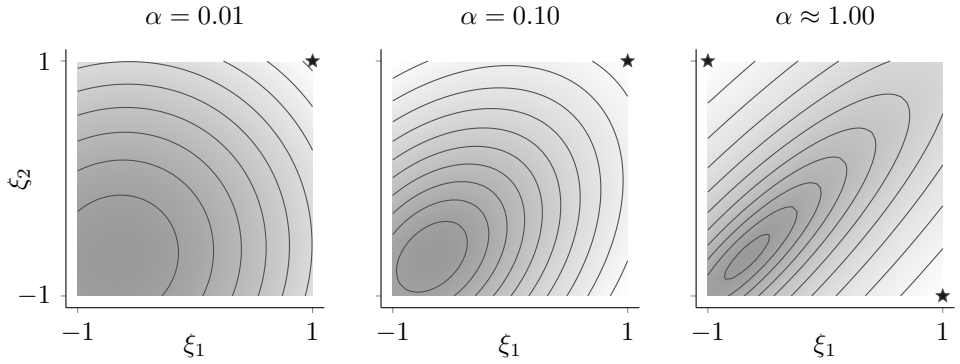

Figure 3: Robust expected information gain for a two-dimensional linear regression problem with a correlated prior, as a function of $\alpha \in (0, 1)$. Contour lines depict the objective landscape over (ξ_1, ξ_2), highlighting how the optimal designs ($\star$) shift with α.

where the worst-case posterior $q^\star(\theta \mid x_{1:N}, \xi_{1:N}) = \mathrm{N}(\theta; \mu_\theta^\star, \Sigma_\theta^\star)$ is:

$$\Sigma_\theta^\star = \left[\Sigma_0^{-1} + \frac{\alpha}{\sigma^2} H(\xi_{1:N})^\top H(\xi_{1:N})\right]^{-1}, \quad \mu_\theta^\star = \Sigma_\theta^\star \left[\Sigma_0^{-1}\mu_0 + \frac{\alpha}{\sigma^2} H(\xi_{1:N})^\top x_{1:N}\right].$$

Sibson's α-MI. Finally, for this conjugate setting, Sibson's α-mutual information admits a closed-form expression. Using the definition

$$I_\alpha^S(\theta; x_{1:N})(\xi_{1:N}) = \frac{\alpha}{\alpha - 1} \log \int \left[\int p(\theta)\, p(x_{1:N} \mid \xi_{1:N}, \theta)^\alpha \, \mathrm{d}\theta\right]^{1/\alpha} \mathrm{d}x_{1:N},$$

which corresponds to the log-normalizer of $p_\alpha(x_{1:N} \mid \xi_{1:N})$:

$$\begin{aligned} I_\alpha^S(\theta; x_{1:N})(\xi_{1:N}) &= \frac{1}{2} \log \left| H(\xi_{1:N})\, \Sigma_0\, H(\xi_{1:N})^\top + \frac{\sigma^2}{\alpha} I_N \right| + \text{const} \\ &= \frac{1}{2} \log \left| \frac{\alpha}{\sigma^2} H(\xi_{1:N})\, \Sigma_0\, H(\xi_{1:N})^\top + I_N \right|. \end{aligned}$$

As $\alpha \to 1$, the robust expected information gain recovers Shannon's mutual information, which is the standard result for Bayesian linear regression with N measurements. As $\alpha \to 0$, the term $\frac{\sigma^2}{\alpha} I_N$ dominates, and all designs become equally uninformative, consistent with Proposition 2.

Rényi Divergence. The Rényi divergence between two Gaussian distributions is useful because it provides tractable computation of the realized robust information gain conditioned on a design and measured outcome. We derive the closed-form for orders $\alpha \in (0, 1)$. Consider two Gaussian distributions $q(\theta) = \mathrm{N}(\theta; \mu_q, \Sigma_q)$ and $p(\theta) = \mathrm{N}(\theta; \mu_p, \Sigma_p)$. The divergence is defined as:

$$\mathbb{D}_\alpha[q \,\|\, p] = \frac{1}{\alpha - 1} \log \int q(\theta)^\alpha \, p(\theta)^{1-\alpha} \, \mathrm{d}\theta.$$

Substituting the Gaussian density functions, the integrand becomes a product of exponential quadratic forms. The exponent corresponds to a new quadratic form defined by the geometric mixture of the parameters. Let Σ_α and μ_α denote the covariance and mean of this geometric mixture, defined by the weighted sum of precision matrices:

$$\Sigma_\alpha = \Big[\alpha\Sigma_q^{-1} + (1-\alpha)\Sigma_p^{-1}\Big]^{-1}, \quad \mu_\alpha = \Sigma_\alpha\Big[\alpha\Sigma_q^{-1}\mu_q + (1-\alpha)\,\Sigma_p^{-1}\mu_p\Big].$$

Evaluating the Gaussian integral yields the log-integral term:

$$\log \int q(\theta)^\alpha \, p(\theta)^{1-\alpha} \, \mathrm{d}\theta = \frac{1}{2} \log \frac{|\Sigma_\alpha|}{|\Sigma_q|^\alpha |\Sigma_p|^{1-\alpha}} - \frac{1}{2} Z_\alpha(\mu_q, \mu_p),$$

where Z_α is defined as follows:

$$Z_\alpha(\mu_q, \mu_p) \coloneqq \alpha\, \mu_q^\top \Sigma_q^{-1} \mu_q + (1-\alpha)\, \mu_p^\top \Sigma_p^{-1} \mu_p - \mu_\alpha^\top \Sigma_\alpha^{-1} \mu_\alpha.$$

Finally, scaling by $1/(\alpha - 1)$ gives the divergence:

$$\mathbb{D}_\alpha[q \,\|\, p] = \frac{1}{2(1-\alpha)} Z_\alpha(\mu_q, \mu_p) + \frac{1}{2(\alpha - 1)} \log \frac{|\Sigma_\alpha|}{|\Sigma_q|^\alpha |\Sigma_p|^{1-\alpha}}.$$

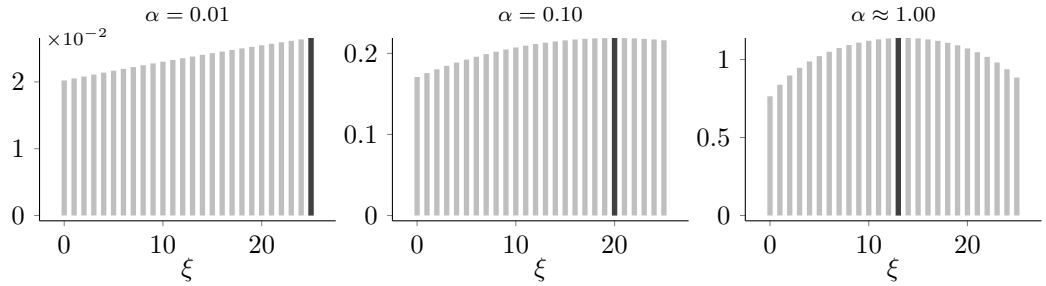

Figure 4: Robust expected information gain as a function of $\alpha \in (0, 1)$ for an A/B testing problem with 25 participants. The optimal allocation, highlighted in dark gray, shifts as α varies.

F.2 A/B Testing

To illustrate the robust experimental design framework in a setting with discrete measurements, we derive expressions for the worst-case distribution and robust expected information gain in the conjugate Beta-Binomial A/B testing setting.

Consider a standard A/B testing scenario where we wish to infer the conversion rates of two independent variants. Let $\theta = (\theta_a, \theta_b)^\top \in [0, 1] \times [0, 1]$ denote the unknown parameter vector, where θ_a and θ_b are the conversion probabilities for groups A and B, respectively. We equip θ with a product of independent Beta priors:

$$p(\theta) = \mathrm{Beta}(\theta_a; \delta_a, \gamma_a) \times \mathrm{Beta}(\theta_b; \delta_b, \gamma_b),$$

where $\delta_a, \gamma_a, \delta_b, \gamma_b > 0$ are the prior hyperparameters. A design $\xi = (n_a, n_b) \in \Xi$ specifies the sample sizes allocated to each group, subject to a total budget constraint $n_a + n_b = N_x$. The measurements $x = (x_a, x_b) \in \{0, \ldots, n_a\} \times \{0, \ldots, n_b\}$ represent the number of conversions, or successes, in each group. Under the assumption of independence, the likelihood factorizes as a product of Binomial distributions:

$$\begin{aligned} p(x \mid \theta, \xi) &= \mathrm{Bin}(x_a; n_a, \theta_a) \times \mathrm{Bin}(x_b; n_b, \theta_b) \\ &= \left[\binom{n_a}{x_a} \theta_a^{x_a} (1 - \theta_a)^{n_a - x_a} \right] \left[\binom{n_b}{x_b} \theta_b^{x_b} (1 - \theta_b)^{n_b - x_b} \right]. \end{aligned}$$

Marginal and Posterior. Under these assumptions, we can derive the marginal likelihood of the measurements $p(x \mid \xi)$ by integrating out the parameters θ. This results in the Beta-Binomial distribution. For a single group $k \in \{a, b\}$, the marginal probability of observing x_k successes in n_k trials is:

$$\begin{aligned} p(x_k \mid n_k) &= \int_0^1 \mathrm{Bin}(x_k; n_k, \theta_k) \, \mathrm{Beta}(\theta_k; \delta_k, \gamma_k) \, \mathrm{d}\theta_k \\ &= \frac{1}{B(\delta_k, \gamma_k)} \binom{n_k}{x_k} \int_0^1 \theta_k^{x_k + \delta_k - 1} (1 - \theta_k)^{n_k - x_k + \gamma_k - 1} \, \mathrm{d}\theta_k \\ &= \frac{1}{B(\delta_k, \gamma_k)} \binom{n_k}{x_k} B(\delta_k + x_k, \gamma_k + n_k - x_k), \end{aligned}$$

where $B(\cdot)$ is the Beta function and the joint marginal likelihood is simply the product $p(x \mid \xi) = p(x_a \mid n_a) \, p(x_b \mid n_b)$.

Furthermore, under the assumptions of conjugacy, we can compute the Bayes posterior $p(\theta \mid x, \xi)$ in closed-form as factorized independent Beta distributions for each group. For $k \in \{a, b\}$:

$$\begin{aligned} p(\theta_k \mid x_k, n_k) &\propto \mathrm{Bin}(x_k; n_k, \theta_k) \, \mathrm{Beta}(\theta_k; \delta_k, \gamma_k) \\ &\propto \theta_k^{x_k} (1 - \theta_k)^{n_k - x_k} \, \theta_k^{\delta_k - 1} (1 - \theta_k)^{\gamma_k - 1} \\ &= \theta_k^{\delta_k + x_k - 1} \, (1 - \theta_k)^{\gamma_k + n_k - x_k - 1} \\ &= \mathrm{Beta}(\theta_k; \delta_k + x_k, \gamma_k + n_k - x_k). \end{aligned}$$

Worst-Case Distribution. The worst-case joint distribution $q^\star(\theta, x \mid \xi)$ is given by:

$$q^\star(\theta, x \mid \xi) \propto \Big[p(\theta) \, p_\alpha(x \mid \xi) \Big]^{1 - \alpha} \Big[p(\theta, x \mid \xi) \Big]^{\alpha},$$

where the α-tilted marginal $p_\alpha(x \mid \xi)$ is defined as:

$$p_\alpha(x \mid \xi) \propto \left[\mathbb{E}_{p(\theta)}\left[p(x \mid \theta, \xi)^\alpha\right]\right]^{1/\alpha}.$$

For this conjugate problem, we compute the expectation of the α-powered likelihood. Due to independence, the expectation factorizes over the categories for θ_a and θ_b:

$$\mathbb{E}_{p(\theta)}\left[p(x \mid \theta, \xi)^\alpha\right] = \prod_{k \in \{a,b\}} \mathbb{E}_{p(\theta_k)}\left[p(x_k \mid \theta_k, n_k)^\alpha\right].$$

Focusing on a single group with parameters (θ_k, n_k, x_k) where $k \in \{a, b\}$:

$$p(x_k \mid \theta_k, n_k)^\alpha = \binom{n_k}{x_k}^\alpha \theta_k^{\alpha x_k} (1-\theta_k)^{\alpha(n_k - x_k)}.$$

The expectation with respect to the Beta prior is:

$$\begin{aligned}
\mathbb{E}_{p(\theta_k)}\left[p(x_k \mid \theta_k, n_k)^\alpha\right] &= \int_0^1 \frac{\theta_k^{\delta_k - 1}(1-\theta_k)^{\gamma_k - 1}}{B(\delta_k, \gamma_k)} \binom{n_k}{x_k}^\alpha \theta_k^{\alpha x_k}(1-\theta_k)^{\alpha(n_k - x_k)} \,\mathrm{d}\theta_k \\
&= \frac{1}{B(\delta_k, \gamma_k)} \binom{n_k}{x_k}^\alpha \int_0^1 \theta_k^{\delta_k + \alpha x_k - 1}(1-\theta_k)^{\gamma_k + \alpha(n_k - x_k) - 1} \,\mathrm{d}\theta_k \\
&= \frac{1}{B(\delta_k, \gamma_k)} \binom{n_k}{x_k}^\alpha B(\delta_k + \alpha\, x_k, \gamma_k + \alpha(n_k - x_k)).
\end{aligned}$$

Let us define the auxiliary term $Z_{\alpha,k}(x_k, n_k)$ as this expectation:

$$Z_{\alpha,k}(x_k, n_k) \coloneqq \frac{1}{B(\delta_k, \gamma_k)} \binom{n_k}{x_k}^\alpha B(\delta_k + \alpha\, x_k, \gamma_k + \alpha(n_k - x_k)).$$

The α-tilted marginal $p_\alpha(x \mid \xi)$ then takes the form:

$$p_\alpha(x \mid \xi) \propto \left[Z_{\alpha,a}(x_a, n_a) Z_{\alpha,b}(x_b, n_b)\right]^{1/\alpha}.$$

To normalize this distribution, we compute the partition function:

$$\mathcal{Z}_\alpha(\xi) = \left[\sum_{x_a=0}^{n_a} Z_{\alpha,a}(x_a, n_a)^{1/\alpha}\right]\left[\sum_{x_b=0}^{n_b} Z_{\alpha,b}(x_b, n_b)^{1/\alpha}\right].$$

Consequently, the worst-case joint distribution $q^\star(\theta, x \mid \xi)$ implies a worst-case posterior $q^\star(\theta \mid x, \xi)$ which is the α-tilted posterior. For the Beta-Binomial model, this results in a Beta distribution with updated parameters:

$$q^\star(\theta_k \mid x_k, n_k) = \mathrm{Beta}(\theta_k; \delta_k + \alpha\, x_k, \gamma_k + \alpha(n_k - x_k)).$$

This confirms that under model misspecification, $\alpha < 1$, the experimenter updates beliefs less aggressively than in the standard Bayesian setting.

Sibson's α-MI. Sibson's α-mutual information can be written as:

$$I_\alpha^S(\theta; x)(\xi) = \frac{\alpha}{\alpha - 1} \log \sum_x \left[\mathbb{E}_{p(\theta)}\left[p(x \mid \theta, \xi)^\alpha\right]\right]^{1/\alpha}.$$

Substituting our derived terms, this becomes $I_\alpha^S(\xi) = \alpha/(\alpha-1) \log \mathcal{Z}_\alpha(\xi)$. Due to the independence of the groups, the information gain is additive $I_\alpha^S(\xi) = I_\alpha^S(n_a) + I_\alpha^S(n_b)$, where the contribution of a single group with allocation n_k is:

$$I_\alpha^S(n_k) = \frac{\alpha}{\alpha - 1} \log \left\{ \sum_{x_k=0}^{n_k} \left[\binom{n_k}{x_k}^\alpha \frac{B(\delta_k + \alpha\, x_k, \gamma_k + \alpha(n_k - x_k))}{B(\delta_k, \gamma_k)}\right]^{1/\alpha} \right\}.$$

As $\alpha \to 1$, we recover the standard expected information gain for the Beta-Binomial model, which corresponds to Shannon's mutual information. As $\alpha \to 0$, the robust information gain approaches zero.

Rényi Divergence. Similar to the Gaussian case, the Rényi divergence between two Beta distributions admits a closed-form expression, which allows for tractable computation of the realized robust

information gain given a design and outcome. Consider two Beta distributions $q(\theta) = \mathrm{Beta}(\theta; \delta_q, \gamma_q)$ and $p(\theta) = \mathrm{Beta}(\theta; \delta_p, \gamma_p)$. The divergence of order α is:

$$\mathbb{D}_\alpha[q \,\|\, p] = \frac{1}{\alpha - 1} \log \int_0^1 q(\theta)^\alpha \, p(\theta)^{1-\alpha} \, \mathrm{d}\theta.$$

Substituting the Beta density functions, the integrand becomes a product of power functions. This results in an unnormalized Beta density with a geometric mixture of the original parameters. Let δ_α and γ_α denote the effective parameters of this mixture, defined by the linear interpolation of the natural parameters:

$$\delta_\alpha = \alpha(\delta_q - 1) + (1 - \alpha)(\delta_p - 1) + 1, \quad \gamma_\alpha = \alpha(\gamma_q - 1) + (1 - \alpha)(\gamma_p - 1) + 1.$$

The integral can then be evaluated analytically using the Beta function $B(\cdot)$:

$$\int_0^1 q(\theta)^\alpha \, p(\theta)^{1-\alpha} \, \mathrm{d}\theta = \frac{B(\delta_\alpha, \gamma_\alpha)}{B(\delta_q, \gamma_q)^\alpha \, B(\delta_p, \gamma_p)^{1-\alpha}}.$$

Taking the logarithm and scaling by $1/(\alpha - 1)$ yields the closed-form divergence:

$$\mathbb{D}_\alpha[q \,\|\, p] = \frac{1}{\alpha - 1} \log B(\delta_\alpha, \gamma_\alpha) - \frac{\alpha}{\alpha - 1} \log B(\delta_q, \gamma_q) - \frac{1 - \alpha}{\alpha - 1} \log B(\delta_p, \gamma_p).$$

G Connection to Barlas et al. [2025]

In this section, we contrast Sibson's α-mutual information with the power-likelihood version of the *Gibbs expected information gain* (GEIG) presented by Barlas et al. [2025]. For a power likelihood with exponent $\alpha > 0$, the generalized likelihood $p(x \mid \theta, \xi)^\alpha$ is not a normalized density in x. Rather, it is the unnormalized likelihood factor used in the Gibbs posterior

$$p_\alpha(\theta \mid x, \xi) = \frac{p(\theta) \, p(x \mid \theta, \xi)^\alpha}{\int p(\vartheta) \, p(x \mid \vartheta, \xi)^\alpha \, \mathrm{d}\vartheta},$$

and the denominator is the marginal generalized likelihood

$$m_\alpha(x \mid \xi) \coloneqq \int p(\theta) \, p(x \mid \theta, \xi)^\alpha \, \mathrm{d}\theta.$$

The marginal $m_\alpha(x \mid \xi)$ need not integrate to one over x either. Under the *pseudo-information* construction of Barlas et al. [2025],[1] the power-likelihood Gibbs expected information gain is

$$I^G_\alpha(\xi) \coloneqq \iint p(\theta) \, p(x \mid \theta, \xi)^\alpha \Big[\alpha \log p(x \mid \theta, \xi) - \log m_\alpha(x \mid \xi)\Big] \, \mathrm{d}\theta \, \mathrm{d}x.$$

The powered likelihood thus appears in three places. It defines the unnormalized measure of the outer *pseudo-expectation*, it enters the logarithm through $\alpha \log p(x \mid \theta, \xi)$, and it enters the marginal generalized likelihood $m_\alpha(x \mid \xi)$. The omission of the normalizing constants has direct consequences for the resulting information measure. To make this explicit, suppose we instead normalize the α-powered likelihood over x and define the α-powered density

$$f_\alpha(x \mid \theta, \xi) \coloneqq \frac{p(x \mid \theta, \xi)^\alpha}{\int p(x \mid \theta, \xi)^\alpha \, \mathrm{d}x} = \frac{p(x \mid \theta, \xi)^\alpha}{C_\alpha(\theta, \xi)}.$$

Replacing $p(x \mid \theta, \xi)^\alpha$ by $f_\alpha(x \mid \theta, \xi)$ leaves the Gibbs posterior unchanged only if $C_\alpha(\theta, \xi)$ does not depend on θ. Rewriting the marginal generalized likelihood in terms of f_α gives

$$m_\alpha(x \mid \xi) = \int p(\theta) \, C_\alpha(\theta, \xi) \, f_\alpha(x \mid \theta, \xi) \, \mathrm{d}\theta.$$

In general, the normalizer $C_\alpha(\theta, \xi)$ is therefore not an irrelevant constant. If it depends on θ, it changes the relative contribution of different parameter values to the marginal generalized likelihood. If it depends on ξ, it can also change the ranking of designs. Only in special cases where $C_\alpha(\theta, \xi)$ is independent of both θ and ξ can it be dropped without affecting the posterior over θ or the design selected by maximizing the Gibbs expected information gain.

[1] We follow the first version of Barlas et al. [2025] (arXiv v1, November 2025). A later revision (v2, October 2026) normalizes this pseudo-expectation. Under that normalization, the GEIG coincides with Sibson's α-mutual information in the linear regression example below.

Linear Regression. We examine the linear regression case detailed in Appendix F.1. In this case, the likelihood of an outcome is defined as:

$$p(x_{1:N} \mid \theta, \xi_{1:N}) = \mathrm{N}\big(x_{1:N}; H(\xi_{1:N})\,\theta, \sigma^2 I_N\big)$$

$$p(x_{1:N} \mid \theta, \xi_{1:N})^\alpha = C_\alpha\,\mathrm{N}\big(x_{1:N}; H(\xi_{1:N})\,\theta, \frac{\sigma^2}{\alpha} I_N\big), \qquad C_\alpha = (2\pi\sigma^2)^{N(1-\alpha)/2}\alpha^{-N/2}.$$

Hence the marginal is

$$m_\alpha(x_{1:N} \mid \xi_{1:N}) = C_\alpha\,\mathrm{N}\big(x_{1:N}; H(\xi_{1:N})\,\mu_0, H(\xi_{1:N})\,\Sigma_0\,H(\xi_{1:N})^\top + \frac{\sigma^2}{\alpha} I_N\big),$$

and

$$I^G_\alpha(\xi_{1:N}) = C_\alpha\,\frac{1}{2}\log\left|I_N + \frac{\alpha}{\sigma^2} H(\xi_{1:N})\,\Sigma_0\,H(\xi_{1:N})^\top\right| = C_\alpha\,I^S_\alpha(\theta; x_{1:N})(\xi_{1:N}).$$

Thus, when N and σ^2 are fixed, the power-likelihood generalized EIG from Barlas et al. [2025] has the same maximizers as our closed-form Sibson objective for linear regression, but its value is scaled by the unnormalized generalized likelihood mass.

This is a special case where the normalizer C_α is independent of θ and ξ, so the omission of the normalizing constant does not affect the design selected by maximizing the acquisition function. However, this is not a general phenomenon, and in other settings the normalizer can have a significant impact on the resulting information measure and optimal design.

A/B Testing. For the A/B testing problem, detailed in Appendix F.2, the same simplification does not occur. The α-powered Binomial likelihood has a normalizing constant that depends on the allocation, so the effect of using the unnormalized α-powered likelihood cannot, in general, be reduced to a design-independent scale factor.

In this example, the likelihood of an outcome is given by the product of Binomial distributions:

$$p(x \mid \theta, \xi) = \mathrm{Bin}(x_a; n_a, \theta_a) \times \mathrm{Bin}(x_b; n_b, \theta_b).$$

For a single variant, define the normalizer of the α-powered likelihood by

$$C_{k,\alpha}(\theta_k, n_k) \coloneqq \sum_{u=0}^{n_k} \mathrm{Bin}(u; n_k, \theta_k)^\alpha = \sum_{u=0}^{n_k} \binom{n_k}{u}^\alpha \theta_k^{\alpha u}\,(1-\theta_k)^{\alpha(n_k-u)}.$$

For $\alpha = 1$, this normalizer is equal to one. For $\alpha < 1$, it is generally a non-constant function of both θ_k and n_k. Thus the normalized α-powered conditional density would be

$$f_{k,\alpha}(x_k \mid \theta_k, n_k) \coloneqq \frac{\mathrm{Bin}(x_k; n_k, \theta_k)^\alpha}{C_{k,\alpha}(\theta_k, n_k)}.$$

The joint powered likelihood can therefore be written as $p(x \mid \theta, \xi)^\alpha = C_\alpha(\theta, \xi)\,f_\alpha(x \mid \theta, \xi)$, where

$$C_\alpha(\theta, \xi) \coloneqq \prod_{k \in \{a,b\}} C_{k,\alpha}(\theta_k, n_k), \qquad f_\alpha(x \mid \theta, \xi) \coloneqq \prod_{k \in \{a,b\}} f_{k,\alpha}(x_k \mid \theta_k, n_k).$$

Remember, the power-likelihood Gibbs EIG is written as

$$I^G_\alpha(\xi) = \sum_x \int p(\theta)\,C_\alpha(\theta, \xi)\,f_\alpha(x \mid \theta, \xi)\Big[\log C_\alpha(\theta, \xi) + \log f_\alpha(x \mid \theta, \xi) - \log m_\alpha(x \mid \xi)\Big]\,\mathrm{d}\theta.$$

By contrast, Sibson's α-mutual information for the same problem is

$$I^S_\alpha(\theta; x)(n_a, n_b) = \frac{\alpha}{\alpha - 1}\log \sum_{x_a=0}^{n_a}\sum_{x_b=0}^{n_b} m_\alpha(x \mid \xi)^{1/\alpha}.$$

For $\alpha < 1$, these two objectives are not equivalent and can induce different optimal designs. We demonstrate this effect empirically across a geometric range of α values in Table 2.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The claims made are backed up by both the theory we develop and the numerical experiments we report.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: We discuss the limitations of our method in the final section of the main text.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification: We provide full derivations with the necessary assumptions for all lemmas and propositions in the appendix.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: The contribution of this work is conceptual and analytical. All quantities entering the numerical evaluations are stated explicitly, either in closed form or as estimators.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [No]

Justification: The contribution of this work is conceptual and analytical. No external data is used.

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

Answer: [Yes]

Justification: The models, priors, sample sizes, and the grids of α and λ are specified in Section 7 and Appendix F.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification: Wherever needed, we report the mean along with the 10th and 90th percentiles of statistical quantities. Our evaluation also includes histograms that depict the empirical distribution of evaluated metrics.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.

- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: As stated in Appendix F, the compute and memory requirements to reproduce the evaluation are negligible: all experiments run on a standard CPU and complete within minutes.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The research complies with the NeurIPS Code of Ethics.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [No]

    Justification: This work presents algorithmic and theoretical contributions to the field of experimental design without addressing specific applications or deployment scenarios, hence we do not anticipate any direct positive or negative societal consequences.

Guidelines:

- The answer [N/A] means that there is no societal impact of the work performed.
- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: We do not see a potential risk of misuse for our work.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: We have provided proper accreditation wherever applicable.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.

- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: No new assets are released. See earlier comments on open access to data and code.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The work does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: This work does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.

- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: This work does not involve LLMs.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.