Admissibility of Completely Randomized Trials: A Large-Deviation Approach

Guido Imbens
Stanford University
imbens@stanford.edu

Chao Qin
Stanford University
chaoqin@stanford.edu

Stefan Wager
Stanford University
swager@stanford.edu

1 Introduction

Randomized control trials (RCTs) are considered a gold-standard method for causal inference and data-driven decision making in many fields (Fisher 1925, Imbens and Rubin 2015). The traditional approach to RCT designs involves pre-specifying the treatment assignment mechanism before any data is collected; for example, in completely randomized trials (CRTs), the actual treatment assignment is chosen uniformly at random from all possible assignments with the same marginal treatment fractions (Fisher 1925). Recently, however, there has been growing concern that in settings where an analyst wants to learn about multiple treatment arms and has the option to run an adaptive experiment, standard designs such as CRTs may be inefficient relative to adaptive designs that can use data collected early in the trial to better target their experimentation budget. Such considerations can arise, for example, in online marketing (Chapelle and Li 2011), interface design (Qin and Russo 2023), job-search assistance (Caria et al. 2024), or vaccine trials (Wu and Wager 2022). It is by now clear that, in some application areas, adaptive trials can vastly outperform non-adaptive designs (e.g., Chapelle and Li 2011). What's less clear, however, is whether successful deployment of adaptive experimental designs fundamentally relies on the use of problem-specific domain knowledge (in which case basic CRTs would remain attractive as a robust, domain-agnostic baseline method), or whether there exist adaptive designs that uniformly dominate CRTs (Qin 2022).

We answer this question affirmatively in the context of the well-known best-arm identification problem (Audibert et al. 2010), known in the simulation literature as the ranking and selection problem (Glynn and Juneja 2004, Hong et al. 2021). The goal is to deploy a treatment arm at the end of the experiment with high confidence that it is the best, or achieves low degradation in welfare. We show that—for this task and using a large-deviation error metric (described further in Definition 1 below)—there exist simple adaptive designs that dominate standard CRTs on an instance-by-instance level whenever there are at least $K \geq 3$ treatment arms to choose from. This implies that standard CRTs are not admissible among potentially adaptive randomized experiments in a sense analogous to that of Wald (1949). Our result provides an affirmative answer to the second open problem posed in Qin (2022). Our dominance results are achieved within a class of "batched arm elimination" (BAE) designs, presented in Section 2. BAE designs sequentially discard the worst-performing arms from the experiment at pre-specified checkpoints. When there are only $K = 2$ treatment arms, BAE designs reduce to standard CRTs. However, when $K \geq 3$ arms are available, we show that simple BAE designs can uniformly outperform CRTs. Here is an example of such BAE designs: given T experimental units:

1. Run a completely randomized trial with all K arms on $\left\lceil \frac{K}{2(K-1)} T \right\rceil$ units.
2. Discard the worst-performing arm after the first batch.
3. Run a completely randomized trial with the remaining $(K-1)$ arms on all remaining units.
4. Select the best of these arms based on aggregate empirical performance across both batches.

These dominance results are derived from an exact characterization of the large-deviation behavior of BAE designs, which is presented in Section 3. In Appendix B, we demonstrate our proposed design on a semi-synthetic experiment calibrated to a randomized trial by Karlan and List (2007).

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: MLxOR: Mathematical Foundations and Operational Integration of Machine Learning for Uncertainty-Aware Decision-Making.

Problem formulation. We frame our analysis in terms of a standard i.i.d. sampling model for multi-armed experimentation (Lattimore and Szepesvári 2020). An experimenter conducts an adaptive experiment to identify the best treatment arm among K arms to deploy, following sequentially assigning these K treatments to T experimental units. The potential outcome of assigning treatment $i \in [K] \triangleq \{1, \ldots, K\}$ to experimental unit $t \in [T] \triangleq \{1, \ldots, T\}$ is a scalar random variable $Y_{t,i}$, where larger values indicate more desirable outcomes. We assume that for each treatment i, the potential outcomes $(Y_{t,i})_{t\in[T]}$ are drawn i.i.d. from a distribution $P(\cdot \mid \theta_i)$ with an unknown scalar parameter θ_i. If these parameters were known, the experimenter would deploy an arm with the highest expected outcome, given by $\max_{i\in[K]} \int y \cdot P(\mathrm{d}y \mid \theta_i)$. Let $\boldsymbol{\theta} \triangleq (\theta_1, \ldots, \theta_K)$ denote the vector of unknown parameters, which we refer to as the *problem instance*. Since these parameters are unknown, the experimenter interacts sequentially with experimental units to learn which treatment arm is best. For the t-th experimental unit, the experimenter selects a treatment arm $I_t \in [K]$ based on the history of previously assigned treatments and observed outcomes, denoted by $H_{t-1} \triangleq \{I_1, Y_{1,I_1}, \ldots, I_{t-1}, Y_{t-1,I_{t-1}}\}$. Importantly, only the outcome of the chosen treatment, Y_{t,I_t}, is observed; outcomes of the unselected treatments remain unknown. The experimenter's goal is to identify and deploy the best treatment arm among K arms with high confidence by the end of the experiment. The experimenter needs to design a policy π, which is a (potentially randomized) decision rule that governs both the sequential allocation of treatment arms to T experimental units and the final deployment of a treatment arm; specifically, it consists of: an *allocation rule* that sequentially assigns treatment arms based on observed outcomes, and a *deployment rule* that selects a treatment arm for deployment after all T units have been treated. Formally, the allocation rule is a function that maps the history of past allocations and outcomes, denoted by H_{t-1}, and the sample size T to the treatment assignment I_t for the t-th experimental unit. Additionally, after all T units have been treated, the deployment rule maps the sample size T and the full history H_T to the final deployed arm $\hat{I}_T$. We denote the class of all such policies by Π. The experimenter's objective is to minimize the *post-experiment utilitarian regret* of the arm deployed under policy π—also known as *simple regret*, as termed by Audibert et al. (2010):

$$\mathfrak{R}^{\pi}_{\boldsymbol{\theta},T} \triangleq \max_{i\in[K]} \theta_i - \mathbb{E}^{\pi}_{\boldsymbol{\theta},T}\left[\theta_{\hat{I}_T}\right]. \tag{1}$$

A number of authors have shown that, for analytic tractability, it is helpful to study adaptive experiments in an asymptotic regime where errors can be characterized using large-deviation methods (Chernoff 1959, Glynn and Juneja 2004, Kaufmann et al. 2016, Russo 2020). Here, we also leverage such asymptotics, under which different exploration policies can be usefully compared in terms of the efficiency exponent given below.

Definition 1 (Efficiency exponent)**.** *The efficiency exponent of a policy π for instance $\boldsymbol{\theta}$ is*

$$\mathfrak{e}^{\pi}_{\boldsymbol{\theta}} \triangleq \liminf_{T\to\infty} \; -\tfrac{1}{T} \ln\left(\mathfrak{R}^{\pi}_{\boldsymbol{\theta},T}\right).$$

We refer to a policy as admissible if there exists no other policy that beats it on an instance-by-instance level (with strict inequality for some instance).

Definition 2 (Large-deviation admissible design)**.** *Given a set of candidate instances Θ, a policy π is large-deviation admissible is there is no policy $\tilde{\pi} \in \Pi$ such that*

$$\forall \boldsymbol{\theta} \in \Theta, \quad \mathfrak{e}^{\tilde{\pi}}_{\boldsymbol{\theta}} \geq \mathfrak{e}^{\pi}_{\boldsymbol{\theta}} \quad \textit{and} \quad \exists \boldsymbol{\theta}' \in \Theta, \quad \mathfrak{e}^{\tilde{\pi}}_{\boldsymbol{\theta}'} > \mathfrak{e}^{\pi}_{\boldsymbol{\theta}'}$$

Throughout this paper, for simplicity, we will work under a setting where there is a unique best arm and all arms have Gaussian sampling distributions with the same variance. Under this setting, the efficiency exponent of completely randomized trials is well known. Our challenge will be to find a policy that achieves a higher efficiency exponent on an instance-by-instance level.

Assumption 1. *Let $\sigma^2 > 0$ such that $P(\cdot \mid \theta_i) = \mathcal{N}(\theta_i, \sigma^2)$. Given this class of distributions, we consider the set Θ of problem instances with a unique best arm,*

$$\Theta \triangleq \left\{\boldsymbol{\theta} = (\theta_1, \ldots, \theta_K) \in \mathbb{R}^K \;:\; I^*(\boldsymbol{\theta}) \triangleq \operatorname{argmax}_{i\in[K]} \theta_i \textit{ is a singleton set}\right\}. \tag{2}$$

Proposition 1. *Under Assumption 1, a completely randomized trial that (non-adaptively) uniformly allocates treatment across K available arms achieves an efficiency exponent:*

$$\mathfrak{e}^{\texttt{Unif}}_{\boldsymbol{\theta}} = \frac{\Delta_{\min}(\boldsymbol{\theta})^2}{4K\sigma^2} \quad \textit{where} \quad \Delta_{\min}(\boldsymbol{\theta}) \triangleq \theta_{I^*} - \max_{i\neq I^*} \theta_i. \tag{3}$$

Algorithm 1 Batched arm elimination

Input: Sample size T, and batch weights $(\beta_K, \beta_{K-1}, \dots, \beta_2) \in \Sigma_{K-1}$
Initialize: Number of remaining arms $n \leftarrow K$, and candidate set $C \leftarrow [K]$
for $t = 1, \dots, T$ **do**
    Assign $I_t \in \operatorname{argmin}_{i \in C} N_{t-1,i}$ to the t-th individual
    Update $\{N_{t,i}\}_{i \in C}$ and $\{m_{t,i}\}_{i \in C}$
    if $r \geq 2$ and $t = (\beta_K + \cdots + \beta_r)T$ **then**
        Remove $\ell_n \in \operatorname{argmin}_{i \in C} m_{t,i}$ from C, i.e., $C \leftarrow C \setminus \{\ell_n\}$, and $n \leftarrow n - 1$
    end if
end for
Output: The only remaining arm in C

2 Batched arm elimination

Recall that we seek to design an adaptive policy that can choose a good arm among K options using T datapoints. Batched arm elimination (BAE) initializes the candidate arm set as $[K]$ and divides the whole sample of T experimental units into $K-1$ batches. For each batch, BAE experiments on the treatment arms in the candidate set in a round-robin manner, and discards at the end of the batch an arm from the candidate set with the lowest empirical mean. More specifically, BAE takes as inputs the sample size T and $K-1$ instance-agnostic batch weights $\beta_K, \beta_{K-1}, \dots, \beta_2$, where the sample size $\beta_n \cdot T$ corresponds to the batch with the candidate set of n arms. BAE begins by experimenting on all arms in a round-robin fashion until K arms have been allocated to a total of $\beta_K \cdot T$ experimental units, after which the arm with the lowest sample mean is eliminated. It then proceeds to experiment on the remaining $K-1$ arms using another $\beta_{K-1} \cdot T$ units. This process continues iteratively, reducing the number of arms in each batch by one, until only one arm remains.

We use the following notation throughout. Let Σ_{K-1} be the $(K-2)$-dimensional simplex with $K-1$ entries. Given this, $(\beta_K, \beta_{K-1}, \dots, \beta_2) \in \Sigma_{K-1}$. For $t \leq T$, the number of experimental units that receive treatment i is denoted by $N_{t,i} \triangleq \sum_{\ell=1}^{t} \mathbb{1}(I_\ell = i)$. When it is positive, we define the empirical mean reward as $m_{t,i} \triangleq \frac{\sum_{\ell=1}^{t} \mathbb{1}(I_\ell = i) Y_{\ell,i}}{N_{t,i}}$. When $N_{t,i} = 0$, we let $m_{t,i} = 0$. We give pseudocode for the BAE procedure in Algorithm 1.

We note that BAE is a direct generalization of the "successive rejects" algorithm proposed by Audibert et al. (2010). Successive rejects is BAE with the following batch weights:

$$(\beta_K, \beta_{K-1}, \dots, \beta_3) = \left(1, \tfrac{1}{K}, \dots, \tfrac{1}{4}\right) \cdot \tfrac{1}{\overline{\ln}(K)} \quad \text{and} \quad \beta_2 = 1 - \textstyle\sum_{n=3}^{K} \beta_n.$$

where $\overline{\ln}(K) = \frac{1}{2} + \sum_{i=2}^{K} \frac{1}{i}$. Audibert et al. (2010) did not investigate uniform-dominance results as we do here. Furthermore, the set of BAE procedures we show dominate uniform sampling in fact does not include the original successive rejects algorithm. It can be verified that uniform allocation outperforms successive rejects in instances where all suboptimal arms are identical. Additionally, we note that CRTs are a speical case of BAE designs with batch weights $(1, 0, \dots, 0) \in \Sigma_{K-1}$. In other words, CRTs consists of a single batch that includes the entire sample size.

3 CRT-dominating batched arm elimination

In this section, we demonstrate the universal superiority of BAE designs over CRTs in terms of the efficiency exponent, providing certain sufficient conditions are met. We begin by deriving an exact characterization of the large-deviation behavior of all BAE designs. Our characterization result extends the results of Wang et al. (2023, 2024), which focused on bounded observations and aimed to minimize the probability of misidentifying the best arm, with an emphasis on the successive rejects algorithm (Audibert et al. 2010) and its two variants proposed in Wang et al. (2023). In contrast, we consider unbounded Gaussian observations and aim to minimize the utilitarian regret, studying all BAE designs with any batch weights $(\beta_K, \beta_{K-1}, \dots, \beta_2) \in \Sigma_{K-1}$.

Recall that in BAE designs (Algorithm 1), we use $n \in \{K, \dots, 2\}$ to denote the number of remaining arms. Let $\mathcal{J}_{\boldsymbol{\theta},n}$ be the collection of all sets with cardinality n that includes the best arm $I^* = I^*(\boldsymbol{\theta})$:

$$\mathcal{J}_{\boldsymbol{\theta},n} \triangleq \{J \subseteq [K] : \; |J| = n, \; I^* \in J\}, \tag{4}$$

which depends on $\boldsymbol{\theta}$ since it is defined based on its best arm I^*. Consider such a set $J \in \mathcal{J}_{\boldsymbol{\theta},n}$; we define the set of instances where I^* is the worst-performing arm in this set:

$$\Lambda_{\boldsymbol{\theta},J} \triangleq \left\{ \boldsymbol{\lambda} \in \mathbb{R}^K : \ \lambda_{I^*} \leq \min_{i \in J} \lambda_i \right\}. \tag{5}$$

Building on the previous definitions, we introduce a quantity that captures the minimal information required for the best arm I^* to be eliminated at the end of a batch, starting with n arms, by the remaining $n-1$ arms:

$$\Gamma_{\boldsymbol{\theta},n} \triangleq \min_{J \in \mathcal{J}_{\boldsymbol{\theta},n}} \inf_{\boldsymbol{\lambda} \in \Lambda_{\boldsymbol{\theta},J}} \sum_{i \in J} d(\lambda_i, \theta_i), \tag{6}$$

where $d(\lambda, \theta) = \frac{1}{2\sigma^2}(\lambda - \theta)^2$ is the Kullback–Leibler (KL) divergence between two Gaussian distributions with means λ and θ and common variance σ^2. In addition to the minimal information quantity $\Gamma_{\boldsymbol{\theta},n}$, we compute the proportion (of the sample size T) allocated to the arm eliminated at the end of the batch, starting with n arms:

$$w_n \triangleq \frac{\beta_K}{K} + \frac{\beta_{K-1}}{K-1} + \cdots + \frac{\beta_n}{n}, \tag{7}$$

where the first term arises from the fact that, in the first batch, the proportion β_K (of the sample size T) is uniformly allocated across K arms, and the remaining terms follow the same logic. These quantities introduced above characterize the efficiency exponent of the BAE designs, as established in Lemma 1 below. The proof of Lemma 1 is provided in Appendix D.

Lemma 1. *Under Assumption 1 and for batched arm elimination (Algorithm 1),*

$$\mathfrak{e}_{\boldsymbol{\theta}}^{\texttt{BAE}} \geq \min_{n \in \{K, K-1, \ldots, 2\}} \ w_n \Gamma_{\boldsymbol{\theta},n},$$

where w_n and $\Gamma_{\boldsymbol{\theta},n}$ are defined in (7) *and* (6)*, respectively.*

Our next task is to establish sufficient conditions for BAE designs being universally dominating the CRT. We derive the following lower bound on the information quantity $\Gamma_{\boldsymbol{\theta},n}$ in (6), by analyzing different instance configurations respectively.

Lemma 2 (A lower bound on minimal information $\Gamma_{\boldsymbol{\theta},n}$)**.** *For any $n \in \{K, \ldots, 2\}$,*

$$\Gamma_{\boldsymbol{\theta},n} \geq \frac{n-1}{n} \frac{\Delta_{\min}(\boldsymbol{\theta})^2}{2\sigma^2} \quad \textit{where} \quad \Delta_{\min}(\boldsymbol{\theta}) = \theta_{I^*} - \max_{i \neq I^*} \theta_i.$$

The inequality becomes equality for the instances such that the suboptimal arms are the same, i.e., $\theta_i = \theta_j$ for any $i, j \neq I^*$.

The proof of Lemma 2 is provided in Appendix E. By integrating this lower bound on the minimal information $\Gamma_{\boldsymbol{\theta},n}$ into the efficiency exponent in Lemma 1, we obtain the lower bound for efficiency exponent of BAE designs.

Corollary 1 (A lower bound on BAE's efficiency exponent)**.** *Under Assumption 1 and for batched arm elimination (Algorithm 1),*

$$\mathfrak{e}_{\boldsymbol{\theta}}^{\texttt{BAE}} \geq \frac{\Delta_{\min}(\boldsymbol{\theta})^2}{2\sigma^2} \min_{n \in \{K, \ldots, 2\}} w_n \frac{n-1}{n}.$$

By comparing this lower bound for efficiency exponent of BAE designs with the efficiency exponent of CRTs in Proposition 1, we immediately derive the sufficient conditions for BAE designs to universally outperforms CRTs, based on the batch weights of BAE designs.

Theorem 1 (Sufficient conditions)**.** *Under Assumption 1, for batched arm elimination (Algorithm 1),*

$$\min_{n \in \{K, \ldots, 2\}} \ w_n \frac{n-1}{n} > \frac{1}{2K} \quad \Longrightarrow \quad \mathfrak{e}_{\boldsymbol{\theta}}^{\texttt{BAE}} > \mathfrak{e}_{\boldsymbol{\theta}}^{\texttt{Unif}}, \quad \forall \boldsymbol{\theta} \in \Theta. \tag{8}$$

Recall that w_n, defined in (7), is the proportion (of the sample size T) allocated to the arm eliminated at the end of the batch starting with n arms. As the number of remaining arms n decreases, the proportion w_n increases, while the multiplier $\frac{n-1}{n}$ decreases. The sufficient condition above requires that the proportions allocated to all arms, weighted by their respective multipliers, exceed half the proportion each arm would receive under uniform allocation or CRTs. We present simple BAE designs consisting of only two batches, with batch weights satisfying the sufficient conditions in (8).

Example 1 (Two-batch CRT-dominating BAE)**.** *Consider a two-batch policy that eliminate $s \in [K-2]$ arms after the first batch, and the other $(K-1-s)$ arms at the end of time T. The corresponding batch weights satisfy $\beta_K + \beta_{K-s} = 1$,i.e., $\beta_{K-1} = \beta_{K-2} = \cdots = \beta_{K-s+1} = \beta_{K-s-1} = \cdots = \beta_2 = 0$. Under this, the sufficient conditions for CRT-dominance in* (8) *become*

$$\beta_K > \frac{1}{2} + \frac{1}{2(K-s)}. \tag{9}$$

The example presented in the introduction corresponds to the case $s = 1$, where only one arm is dropped after the first batch. As the total number of arms K grows, the minimal required size of the first batch (i.e., the right-hand side of (9)*) approaches* half *of the total sample size.*

Acknowledgement

A preliminary version of this work appears as a one-page abstract at the 26th ACM Conference on Economics and Computation (EC'25). We are grateful for helpful conversations with Po-An Wang, Junpei Komiyama, Daniel Russo, Whitney Newey, and Pepe Montiel Olea.

References

Kaito Ariu, Masahiro Kato, Junpei Komiyama, Kenichiro McAlinn, and Chao Qin. A comment on "Adaptive treatment assignment in experiments for policy choice". *Econometrica*, 2025.

Jean-Yves Audibert, Sébastien Bubeck, and Rémi Munos. Best arm identification in multi-armed bandits. In *COLT-23th Conference on learning theory-2010*, pages 13–p, 2010.

A Stefano Caria, Grant Gordon, Maximilian Kasy, Simon Quinn, Soha Osman Shami, and Alexander Teytelboym. An adaptive targeted field experiment: Job search assistance for refugees in Jordan. *Journal of the European Economic Association*, 22(2):781–836, 2024.

Olivier Chapelle and Lihong Li. An empirical evaluation of Thompson sampling. In *Advances in Neural Information Processing Systems*, volume 24, pages 2249–2257, 2011.

Herman Chernoff. Sequential design of experiments. *The Annals of Mathematical Statistics*, 30(3):755–770, 1959.

Rémy Degenne. On the existence of a complexity in fixed budget bandit identification. In *Proceedings of Thirty Sixth Conference on Learning Theory*. PMLR, 2023.

R. A. Fisher. The design of experiments. *The Design of Experiments*, 1925.

Aurélien Garivier and Emilie Kaufmann. Optimal best arm identification with fixed confidence. In *Conference on Learning Theory*, pages 998–1027. PMLR, 2016.

Peter Glynn and Sandeep Juneja. A large deviations perspective on ordinal optimization. In *Proceedings of the 2004 Winter Simulation Conference, 2004.*, volume 1. IEEE, 2004.

Vitor Hadad, David A Hirshberg, Ruohan Zhan, Stefan Wager, and Susan Athey. Confidence intervals for policy evaluation in adaptive experiments. *Proceedings of the National Academy of Sciences*, 118(15): e2014602118, 2021.

Keisuke Hirano and Jack R Porter. Asymptotic representations for sequential decisions, adaptive experiments, and batched bandits. *arXiv preprint arXiv:2302.03117*, 2023.

L Jeff Hong, Weiwei Fan, and Jun Luo. Review on ranking and selection: A new perspective. *Frontiers of Engineering Management*, 8(3):321–343, 2021.

Guido W Imbens and Donald B Rubin. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press, 2015.

Dean Karlan and John A List. Does price matter in charitable giving? evidence from a large-scale natural field experiment. *American Economic Review*, 97(5):1774–1793, 2007.

Maximilian Kasy and Anja Sautmann. Adaptive treatment assignment in experiments for policy choice. *Econometrica*, 89(1):113–132, 2021.

Emilie Kaufmann, Olivier Cappé, and Aurélien Garivier. On the complexity of best-arm identification in multi-armed bandit models. *Journal of Machine Learning Research*, 17(1):1–42, 2016.

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

Tor Lattimore and Csaba Szepesvári. *Bandit algorithms*. Cambridge University Press, 2020.

Alexander R Luedtke and Mark J van der Laan. Statistical inference for the mean outcome under a possibly non-unique optimal treatment strategy. *The Annals of Statistics*, 44(2):713, 2016.

Vianney Perchet, Philippe Rigollet, Sylvain Chassang, and Erik Snowberg. Batched bandit problems. *The Annals of Statistics*, 44(2):660–681, 2016.

Chao Qin. Open problem: Optimal best arm identification with fixed-budget. In *Proceedings of Thirty Fifth Conference on Learning Theory*, pages 5650–5654. PMLR, 02–05 Jul 2022.

Chao Qin and Daniel Russo. Adaptive experimentation in the presence of exogenous nonstationary variation. *arXiv preprint arXiv:2202.09036*, 2023.

Chao Qin and Daniel Russo. Optimizing adaptive experiments: A unified approach to regret minimization and best-arm identification. *arXiv preprint arXiv:2402.10592*, 2024.

Daniel Russo. Simple bayesian algorithms for best-arm identification. *Operations Research*, 68(6):1625–1647, 2020.

Abraham Wald. Statistical decision functions. *The Annals of Mathematical Statistics*, pages 165–205, 1949.

Po-An Wang, Ruo-Chun Tzeng, and Alexandre Proutiere. Best arm identification with fixed budget: A large deviation perspective. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.

Po-An Wang, Kaito Ariu, and Alexandre Proutiere. On universally optimal algorithms for A/B testing. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235, 2024.

Han Wu and Stefan Wager. Partial likelihood Thompson sampling. In *Uncertainty in Artificial Intelligence*, pages 2138–2147. PMLR, 2022.

Kelly Zhang, Lucas Janson, and Susan Murphy. Inference for batched bandits. In *Advances in Neural Information Processing Systems*, volume 33, pages 9818–9829, 2020.

A Further literature review

Pure exploration problems consist of an initial adaptive data collection phase followed by a deployment step. Many different problem formulations fall under this broad framework. In this paper, we consider a "fixed-budget" model where the experiment length is given, and we seek the best possible post-experiment guarantees (Audibert et al. 2010, Wang et al. 2023). Another classical model for pure exploration is the "fixed-confidence" model, where the target error rate is taken as given and we seek to guarantee this error rate with the shortest possible expected experiment length (Chernoff 1959, Garivier and Kaufmann 2016). Furthermore, one can use different metrics to quantify errors, including the utilitarian regret of the deployed arm (Kasy and Sautmann 2021). Regardless of the problem formulation, exact finite-sample analyses for these questions present significant analytical challenges; consequently, most high-profile results in this area rely on asymptotics (Chernoff 1959, Glynn and Juneja 2004, Kaufmann et al. 2016, Garivier and Kaufmann 2016, Russo 2020), as do we.

Among pure exploration problems, arguably the fixed-confidence setting has the longest history, dating back to the classical work of Chernoff (1959) on the sequential design of experiments, and optimality under this model is well understood. In particular, Garivier and Kaufmann (2016) demonstrate the existence of universally asymptotically optimal experimental designs under this model: There exist designs that guarantee error rate δ and whose expected stopping time as $\delta \to 0$ has the best possible dependence for every problem instance $\boldsymbol{\theta}$. Qin and Russo (2024) introduce a unified model that bridges the fixed-confidence setting and the classical regret minimization framework of Lai and Robbins (1985), unifying results from both strands of the literature.

However, while the fixed-confidence problem seems a dual to the fixed-budget problem considered here, insights derived under the fixed-confidence model cannot be directly adapted to our setting (Qin 2022). Unlike in the fixed-confidence model, universally asymptotically optimal policies do not exist under the fixed-budget model in full generality: Degenne (2023, Theorem 8) show that no such policy exists for Bernoulli bandits with two arms or for Gaussian bandits with $K > e^{80/3}$ arms, and Degenne (2023) further conjectures that no universally asymptotically optimal policy exists even when $K \geq 3$ arms. Thus, the problem of optimal experimental design under the fixed-budget model is fundamentally more complicated than that under the fixed-confidence model.

Given this context—and especially the non-existence of universally optimal designs in the fixed-budget setting—we fall back on a follow-up question: Are there policies that at least dominate CRTs in terms of their efficiency exponent, or, conversely, are CRTs large-deviation admissible? This question was recently highlighted as an open problem by Qin (2022);[1] and, to the best of our knowledge, remained open until this paper.

We do note that, when there are only $K = 2$ arms, CRTs are difficult to outperform—unlike the case with $K \geq 3$ arms. For two-armed Gaussian bandits, Kaufmann et al. (2016, Theorem 12) prove that Neyman allocation, i.e., CRTs with samples allocated proportionally to arm variances, is universally asymptotically optimal (under an assumption that arm variances are known a-priori). Meanwhile, for two-armed Bernoulli bandits, although Degenne (2023, Theorem 10) show that no universally optimal policy exists, Wang et al. (2024) prove that CRTs remain large-deviation admissible in the sense of Definition 2.

The class of adaptive algorithms we design to beat CRTs under the fixed-budget model is an adaption of the "successive rejects" algorithm of (Audibert et al. 2010), and falls under the general class of batched bandit algorithms (Perchet et al. 2016). One important question left open in this paper is the question of hypothesis testing using our proposed algorithms; for example, it would useful to provide p-values against the null hypothesis that the chosen arm was in fact sub-optimal. Recent advances in

[1] This is the second open problem posed by Qin (2022); the first one was addressed by the results of Ariu et al. (2025), Degenne (2023).

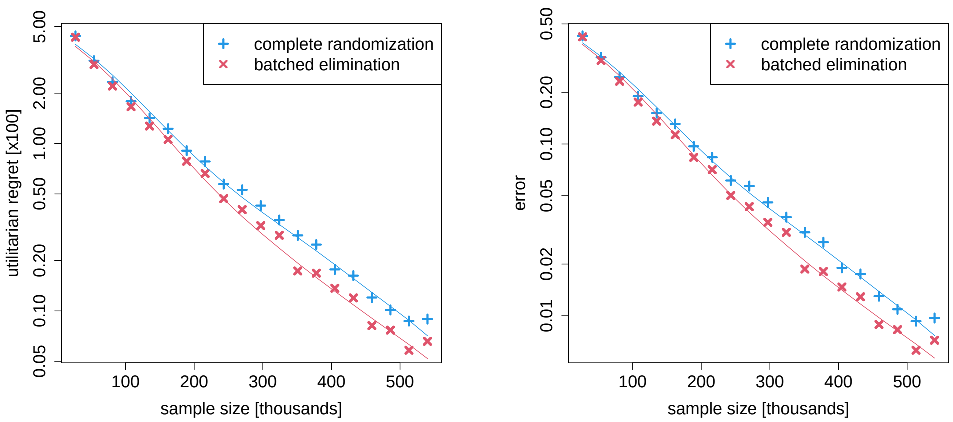

Figure 1: Utilitarian regret (left) and arm-selection error rates (right) for both completely randomized trials and batched arm elimination, as a function of the total sample size T. All results are aggregated across 10,000 simulation replications. Smoothers are added for visualization purposes only.

the literature on inference from adaptively collected data include Hadad et al. (2021), Hirano and Porter (2023), Luedtke and van der Laan (2016) and Zhang et al. (2020).

B Numerical example

Karlan and List (2007) report results on an experiment to test whether matching gifts increase charitable giving. Their experiment includes 4 arms: A control arm, and 3 treatment arms that offer 1:1, 2:1 and 3:1 matches to potential donors (a k:1 match promises that each $1 given will be matched by a $k gift from a different donor). The outcome we are interested in was the total amount given. This is a sparse and skewed outcome: Only 2% of prospective donors give anything (i.e., 98% of the outcomes are 0), and conditionally on donating the mean donation is $44 with a standard deviation of $42. The distribution of the data presents a marked departure of the Gaussianity assumption use in our formal results, and presents us with an opportunity to investigate practical robustness of our algorithm to distributions one might face in real-world applications.

The original experiment of Karlan and List (2007) had a sample size of $n = 50,083$. Here, we start by non-parametrically fitting the data-generating distribution. For each arm we separately tally the fraction of zero outcomes and fit the density of log donation amounts for non-zero donations via kernel density estimation; and combine these to obtain a zero-inflated and skewed density for the outcomes themselves. We then simulate data from this fitted distribution to compare the performance of the following two algorithms for different sample sizes T:

- Completely randomized trial: We simply allocate $T/4$ samples to each arm.
- The variant of batched arm elimination described in Corollary 1 with $s = 1$: We run a completely randomized trial on all arms using $\frac{2}{3}T$ of the data, eliminate the worst-performing arm, and run a completely randomized trial on the remaining 3 arms using the rest of the data.

Throughout, we pick T so that no rounding is required.

Results are shown in Figure 1. We see that both experimental designs perform comparably with smaller sample sizes (in which case they both frequently make errors); however, as the sample size grows, the error of batched arm elimination decays faster than than of the non-adaptive design. Thus, at least in this design, batched arm elimination appears to give a “safe” improvement over non-adaptive experiments: We can benefit from adaptive in high-signal regimes without compromising performance in low-signal ones.

C Proof of Proposition 1

Proof of Proposition 1. Recall that Russo (2020, Proposition 2) establishes that

$$\lim_{T\to\infty} -\frac{1}{T}\ln\left(\mathbb{P}^{\texttt{Unif}}_{\boldsymbol{\theta},T}\left(\hat{I}_T \neq I^*\right)\right) = \frac{\Delta_{\min}(\boldsymbol{\theta})^2}{4K\sigma^2}.$$

Since for any policy π,

$$\Delta_{\min}(\boldsymbol{\theta}) \cdot \mathbb{P}^{\pi}_{\boldsymbol{\theta},T}\left(\hat{I}_T \neq I^*\right) \leq \mathfrak{R}^{\pi}_{\boldsymbol{\theta},T} \leq \Delta_{\max}(\boldsymbol{\theta}) \cdot \mathbb{P}^{\pi}_{\boldsymbol{\theta},T}\left(\hat{I}_T \neq I^*\right),$$

where $\Delta_{\max} \triangleq \theta_{I^*} - \min_{i\neq I^*}\theta_i$, the equality in (3) immediately from Russo (2020, Proposition 2), noting that $\Delta_{\min}(\boldsymbol{\theta}) > 0$ by uniqueness of the best arm and that the optimality gaps $\Delta_{\min}$ and $\Delta_{\max}$ become irrelevant in the asymptotic regime □

D Proof of Lemma 1

For any policy π, we have the following bounds on its utilitarian regret $\mathfrak{R}^{\pi}_{\boldsymbol{\theta},T}$, defined in (1):

$$\Delta_{\min}(\boldsymbol{\theta}) \cdot \mathbb{P}^{\pi}_{\boldsymbol{\theta},T}\left(\hat{I}_T \neq I^*\right) \leq \mathfrak{R}^{\pi}_{\boldsymbol{\theta},T} \leq \Delta_{\max}(\boldsymbol{\theta}) \cdot \mathbb{P}^{\pi}_{\boldsymbol{\theta},T}\left(\hat{I}_T \neq I^*\right),$$

where $\Delta_{\min} = \theta_{I^*} - \max_{i\neq I^*}\theta_i$ and $\Delta_{\max} = \theta_{I^*} - \min_{i\neq I^*}\theta_i$. We note that $\Delta_{\min}(\boldsymbol{\theta}) > 0$ by uniqueness of the best arm (Assumption 1), and that the optimality gaps become irrelevant in the asymptotic regime. Hence, to prove Lemma 1, it suffices to establish the following result.

Lemma 3. *Under Assumption 1 and for batched arm elimination with batch weights $(\beta_K, \beta_{K-1}, \ldots, \beta_2) \in \Sigma_{K-1}$ (Algorithm 1),*

$$\lim_{T\to\infty} -\frac{1}{T}\ln\left(\mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\hat{I}_T \neq I^*\right)\right) \geq \min_{n\in\{K,K-1,\ldots,2\}} w_n\Gamma_{\boldsymbol{\theta},n},$$

where w_n and $\Gamma_{\boldsymbol{\theta},n}$ are defined in (7) *and* (6)*, respectively.*

By the construction of BAE designs (Algorithm 1), we have

$$\mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\hat{I}_T \neq I^*\right) = \sum_{n\in\{K,\ldots,2\}} \mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\ell_n = I^*\right),$$

where ℓ_n is the arm eliminated at the end of the batch starting with n arms. Hence,

$$\mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\hat{I}_T \neq I^*\right) \leq (K-1)\max_{n\in\{K,\ldots,2\}} \mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\ell_n = I^*\right),$$

and thus

$$\begin{aligned}\liminf_{T\to\infty} -\frac{1}{T}\ln\left(\mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\hat{I}_T \neq I^*\right)\right) &\geq \lim_{T\to\infty} -\frac{1}{T}\ln\left(\max_{n\in\{K,\ldots,2\}} \mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\ell_n = I^*\right)\right)\\ &= \liminf_{T\to\infty} \min_{n\in\{K,\ldots,2\}} -\frac{1}{T}\ln\left(\mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\ell_n = I^*\right)\right)\\ &= \min_{n\in\{K,\ldots,2\}} \liminf_{T\to\infty} -\frac{1}{T}\ln\left(\mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\ell_n = I^*\right)\right), \qquad (10)\end{aligned}$$

where the last equality holds since the set of potential values of n is finite.

By (10), establishing the lower bound in Lemma 3 reduces to prove the following result.

Lemma 4. *For $n \in \{K, \ldots, 2\}$,*

$$\liminf_{T\to\infty} -\frac{1}{T}\ln\left(\mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\ell_n = I^*\right)\right) \geq w_n\Gamma_{\boldsymbol{\theta},n}.$$

Proof. We have

$$\mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\ell_n = I^*\right) = \sum_{J\in\mathcal{J}_{\boldsymbol{\theta},n}} \mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\ell_n = I^*, C_n = J\right) \leq |\mathcal{J}_{\boldsymbol{\theta},n}| \cdot \max_{J\in\mathcal{J}_{\boldsymbol{\theta},n}} \mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\ell_n = I^*, C_n = J\right). \qquad (11)$$

Fix $J \in \mathcal{J}_{\boldsymbol{\theta},n}$. Since the arm ℓ_n is eliminated at the end of $T_n \triangleq (\beta_K + \cdots + \beta_n) \cdot T$ timesteps, the event $\{\ell_n = I^*, C_n = J\}$ implies that $\{\boldsymbol{m}_{T_n} \in M, \boldsymbol{p}_{T_n} \in P\}$, where

$$M = \left\{\boldsymbol{\lambda} \in \mathbb{R}^K : \lambda_{I^*} \leq \lambda_i, \forall i \in J\right\} \quad \text{and} \quad P = \{\boldsymbol{p} \in \Sigma_K : p_i = \tilde{p}, \forall i \in J\}.$$

where $\tilde{p} = \frac{\left(\frac{\beta_K}{K} + \cdots + \frac{\beta_n}{n}\right)T}{T_n} = \frac{\frac{\beta_K}{K} + \cdots + \frac{\beta_n}{n}}{\beta_K + \cdots + \beta_n}$. That is,

$$\mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\ell_n = I^*, C_n = J\right) \leq \mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\boldsymbol{m}_{T_n} \in M, \boldsymbol{p}_{T_n} \in P\right),$$

and thus

$$\begin{aligned}
\liminf_{T\to\infty} -\frac{1}{T} \ln\left(\mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\ell_n = I^*, C_n = J\right)\right) &\geq \liminf_{T\to\infty} -\frac{1}{T} \ln\left(\mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\boldsymbol{m}_{T_n} \in M, \boldsymbol{p}_{T_n} \in P\right)\right) \\
&= (\beta_K + \cdots + \beta_n) \liminf_{T\to\infty} -\frac{1}{T_n} \ln\left(\mathbb{P}^{\texttt{BAE}}_{\boldsymbol{\theta},T}\left(\boldsymbol{m}_{T_n} \in M, \boldsymbol{p}_{T_n} \in P\right)\right) \\
&\geq (\beta_K + \cdots + \beta_n) \inf_{\boldsymbol{p}\in \mathrm{cl}(P)} F_{\boldsymbol{\theta},M}(\boldsymbol{p}) \\
&\geq (\beta_K + \cdots + \beta_n)\, \tilde{p} \inf_{\boldsymbol{\lambda}\in \mathrm{cl}(M)} \sum_{i\in J} d(\lambda_i, \theta_i) \\
&= \left(\frac{\beta_K}{K} + \cdots + \frac{\beta_n}{n}\right) \inf_{\boldsymbol{\lambda}\in \mathrm{cl}(M)} \sum_{i\in J} d(\lambda_i, \theta_i) \\
&= w_n \Gamma_{\boldsymbol{\theta},n},
\end{aligned}$$

where the second inequality uses Theorem 2. Applying the inequality in (11) completes the proof. □

D.1 Large-deviation results

The proof of Lemma 4 above relies on the following large-deviation results, which extend Wang et al. (2023, Theorem 1) for Gaussian distributions.

Theorem 2. *Fix $\boldsymbol{\theta} \in \Theta$. For a non-anticipating algorithm, if the empirical allocation sequence $\{\boldsymbol{p}_t\}_{t\geq 1}$ satisfies the large deviation principle (LDP) upper bound with rate function I, then*

1. *$\{\boldsymbol{m}_t\}_{t\geq 1}$ satisfies the LDP lower bound with rate function*

$$\boldsymbol{\lambda} \mapsto \min_{\boldsymbol{p}\in\Sigma_K} \max\{\Psi_{\boldsymbol{\theta}}(\boldsymbol{\lambda}, \boldsymbol{p}), I(\boldsymbol{p})\},$$

where $\Psi_{\boldsymbol{\theta}}(\boldsymbol{\lambda}, \boldsymbol{p}) \triangleq \sum_{i=1}^K p_i \cdot d(\lambda_i, \theta_i)$;

2. *Given $n \in \{2, \ldots, K\}$, consider a set $J \subseteq \mathbb{R}^K$ such that $|J| = n$ and $I^* \in J$. Then for the set*

$$M \triangleq \left\{\boldsymbol{\lambda} \in \mathbb{R}^K : \ \lambda_{I^*} \leq \lambda_i, \ \forall i \in J\right\}$$

and any Borel subset $P \subseteq \Sigma_K$,

$$\liminf_{t\to\infty} -\frac{1}{t} \ln\left(\mathbb{P}_{\boldsymbol{\theta}}\left(\boldsymbol{m}_t \in M, \ \boldsymbol{p}_t \in P\right)\right) \geq \inf_{\boldsymbol{p}\in \mathrm{cl}(P)} \max\left\{F_{\boldsymbol{\theta},M}(\boldsymbol{p}), I(\boldsymbol{p})\right\},$$

where $F_{\boldsymbol{\theta},M}(\boldsymbol{p}) \triangleq \inf_{\boldsymbol{\lambda}\in \mathrm{cl}(M)} \Psi_{\boldsymbol{\theta}}(\boldsymbol{\lambda}, \boldsymbol{p})$, and $\mathrm{cl}(A)$ denotes the closure of the set A.

E Proof of Lemma 2

Fix $n \in \{K, \ldots, 2\}$. For $\mu_1 > \mu_2 \geq \cdots \geq \mu_n$, we define

$$\Psi(\mu_1, \mu_2, \ldots, \mu_n) \triangleq \inf_{\substack{(\lambda_1,\ldots,\lambda_n)\in\mathbb{R}^n: \\ \lambda_1 \leq \min_{i\in[n]} \lambda_i}} \sum_{i=1}^n (\lambda_i - \mu_i)^2. \tag{12}$$

Given this, we can write

$$\Gamma_{\boldsymbol{\theta},n} = \min_{J\in\mathcal{J}_{\boldsymbol{\theta},n}} \inf_{\boldsymbol{\lambda}\in\Lambda_{\boldsymbol{\theta},J}} \sum_{i\in J} d(\lambda_i, \theta_i) = \frac{1}{2\sigma^2} \min_{J\in\mathcal{J}_{\boldsymbol{\theta},n}} \Psi\left(\theta_{I^*}, \theta_{(2)_J}, \ldots, \theta_{(n)_J}\right),$$

where $(i)_J$ denotes the index of the i-th largest element in the set J for $i = 2, \ldots, n$; note that the largest element in J is I^*.

To prove Lemma 2, it suffices to prove the following result:

Lemma 5 (Bounds on $\Psi(\mu_1, \mu_2, \ldots, \mu_n)$)**.** *The value $\Psi(\mu_1, \mu_2, \ldots, \mu_n)$ can be lower bounded as follows,*

$$\Psi(\mu_1, \mu_2, \ldots, \mu_n) \geq \frac{n-1}{n}(\mu_1 - \mu_2)^2,$$

where the inequality becomes equality when $\mu_2 = \cdots = \mu_n$. On the other hand, $\Psi(\mu_1, \mu_2, \ldots, \mu_n)$ can be upper bounded:

$$\Psi(\mu_1, \mu_2, \ldots, \mu_n) < \sum_{i=2}^{n}(\mu_1 - \mu_i)^2.$$

The proof starts with the calculation of $\Psi(\mu_1, \mu_2, \ldots, \mu_n)$, which uses the following result in Wang et al. (2023):

Proposition 2 (Proposition 1 in Wang et al. (2023))**.** *The value $\Psi(\mu_1, \mu_2, \ldots, \mu_n)$ can be calculated as follows,*

$$\Psi(\mu_1, \mu_2, \ldots, \mu_n) = \begin{cases} \sum_{i=1,n} \left(\mu_i - \frac{\mu_1+\mu_n}{2}\right)^2, & \textit{if } \mu_{n-1} \geq \frac{\mu_1+\mu_n}{2}, \\ \sum_{i=1,n-1,n} \left(\mu_i - \frac{\mu_1+\mu_{n-1}+\mu_n}{3}\right)^2, & \textit{if } \mu_{n-1} < \frac{\mu_1+\mu_n}{2},\ \mu_{n-2} \geq \frac{\mu_1+\mu_{n-1}+\mu_n}{3}, \\ \vdots & \vdots \\ \sum_{i=1}^{n} \left(\mu_i - \frac{\sum_{j=1}^{n}\mu_j}{n}\right)^2, & \textit{if } \mu_{n-1} < \frac{\mu_1+\mu_n}{2},\ \ldots,\ \mu_2 < \frac{\mu_1+\mu_3+\cdots+\mu_n}{n-1}. \end{cases}$$

We can simplify the above formulas by introducing $\Delta_i = \mu_1 - \mu_i$ for $i = 1, 2, \ldots, n$.

Corollary 2. *The value $\Psi(\mu_1, \mu_2, \ldots, \mu_n)$ can be rewritten as follows,*

$$\Psi(\mu_1, \mu_2, \ldots, \mu_n) = \begin{cases} \frac{\Delta_n^2}{2}, & \textit{if } \Delta_{n-1} \leq \frac{\Delta_n}{2}, \\ \left(\sum_{i=n-1}^{n}\Delta_i^2\right) - \frac{\left(\sum_{i=n-1}^{n}\Delta_i\right)^2}{3}, & \textit{if } \Delta_{n-1} > \frac{\Delta_n}{2},\ \Delta_{n-2} \leq \frac{\sum_{i=n-1}^{n}\Delta_i}{3}, \\ \vdots & \vdots \\ \left(\sum_{i=2}^{n}\Delta_i^2\right) - \frac{\left(\sum_{i=2}^{n}\Delta_i\right)^2}{n}, & \textit{if } \Delta_{n-1} > \frac{\Delta_n}{2},\ \ldots,\ \Delta_2 > \frac{\sum_{i=3}^{n}\Delta_i}{n-1}. \end{cases}$$

Proof of Corollary 2. If the m-th condition holds where $m \in [n-1]$, we have

$$\begin{aligned} \Psi(\mu_1, \mu_2, \ldots, \mu_n) &= \sum_{i=1,n-m+1,\ldots,n} \left(\mu_i - \frac{\mu_1 + \sum_{j=n-m+1}^{n}\mu_j}{m+1}\right)^2 \\ &= \sum_{i=1,n-m+1,\ldots,n} \left(\Delta_i - \frac{\sum_{j=n-m+1}^{n}\Delta_j}{m+1}\right)^2. \end{aligned}$$

For notational convenience, write $a = \frac{\sum_{j=n-m+1}^{n}\Delta_i}{m+1}$. Then

$$\begin{aligned} \Psi(\mu_1, \mu_2, \ldots, \mu_n) &= a^2 + \sum_{i=n-m+1}^{n}(\Delta_i - a)^2 \\ &= a^2 + ma^2 + \sum_{i=n-m+1}^{n}\Delta_i^2 - 2a\sum_{i=n-m+1}^{n}\Delta_i \\ &= \sum_{i=n-m+1}^{n}\Delta_i^2 - (m+1)a^2 \\ &= \sum_{i=n-m+1}^{n}\Delta_i^2 - \frac{\left(\sum_{i=n-m+1}^{n}\Delta_i\right)^2}{m+1}. \end{aligned}$$

□

Now we are ready to complete the proof of Lemma 5.

Proof of Lemma 5. The upper bound follows directly from the formulas in Corollary 2.

Now we are going to prove the lower bound. If the m-th condition in Corollary 2 holds where $m \in [n-2]$, we have

$$\begin{aligned}\left(\sum_{i=n-m+1}^{n} \Delta_i^2\right) - \frac{\left(\sum_{i=n-m+1}^{n} \Delta_i\right)^2}{m+1} &\geq \frac{\left(\sum_{i=n-m+1}^{n} \Delta_i\right)^2}{m} - \frac{\left(\sum_{i=n-m+1}^{n} \Delta_i\right)^2}{m+1} \\ &= \frac{\left(\sum_{i=n-m+1}^{n} \Delta_i\right)^2}{m(m+1)} \\ &\geq \frac{(m+1)\Delta_{n-m}^2}{m} > \Delta_2^2,\end{aligned}$$

where the second-to-last inequality applied the m-th condition: $\Delta_{n-m} \leq \frac{\sum_{i=n-m+1}^{n} \Delta_i}{m+1}$.

On the other hand, if the last condition in Corollary 2 holds, applying the Cauchy–Schwarz inequality gives

$$\begin{aligned}\left(\sum_{i=2}^{n} \Delta_i^2\right) - \frac{\left(\sum_{i=2}^{n} \Delta_i\right)^2}{n} &\geq \frac{\left(\sum_{i=2}^{n} \Delta_i\right)^2}{n-1} - \frac{\left(\sum_{i=2}^{n} \Delta_i\right)^2}{n} \\ &= \frac{\left(\sum_{i=2}^{n} \Delta_i\right)^2}{(n-1)n} \\ &\geq \frac{n-1}{n}\Delta_2^2,\end{aligned}$$

where the last inequality simply uses $\Delta_i \geq \Delta_2$ for any $i \in \{2, \ldots, n\}$.

Combining the two cases completes the inequality in Lemma 5. The equality holds if and only if $\Delta_2 = \cdots = \Delta_n$. $\square$