

CONTINUOUS TIME ASYMPTOTIC REPRESENTATIONS FOR ADAPTIVE EXPERIMENTS

KARUN ADUSUMILLI[†]

ABSTRACT. This article develops a continuous-time asymptotic framework for analyzing adaptive experiments—settings in which data collection and treatment assignment evolve dynamically in response to incoming information. A key challenge in analyzing fully adaptive experiments, where the assignment policy is updated after each observation, is that the sequence of policy rules often lack a well-defined asymptotic limit. To address this, we focus instead on the empirical allocation process, which captures the fraction of observations assigned to each treatment over time. We show that, under general conditions, any adaptive experiment and its associated empirical allocation process can be approximated by a limit experiment defined by Gaussian diffusions with unknown drifts and a corresponding continuous-time allocation process. This limit representation facilitates the analysis of optimal decision rules by reducing the dimensionality of the state-space and leveraging the tractability of Gaussian diffusions. We apply the framework to derive optimal estimators, analyze in-sample regret for adaptive experiments, and construct e-processes for anytime-valid inference. Notably, we introduce the first definition of any-time and any-experiment valid inference for multi-treatment settings.

This version: January 5, 2026

[†]Department of Economics, University of Pennsylvania.

1. INTRODUCTION

Adaptive experiments are experiments where data is collected and analyzed continuously, allowing for adjustments or decisions to be made on an ongoing basis. Originating from early work by Wald (1947), Arrow et al. (1949), and others, adaptive experimentation has evolved to encompass a wide range of strategies, including bandit experiments, A/B testing, costly sampling, and best-arm identification. These strategies are now widely used across various fields. For instance, tech companies frequently employ bandit algorithms and A/B testing for applications such as web interface optimization, dynamic pricing, and targeted ad placement (see, Bouneffouf and Rish, 2019 for a survey of applications). In clinical trials, multi-stage or group sequential designs (Wassmer and Brannath, 2016) have become standard, allowing early termination of experiments when strong evidence of positive or negative effects emerges. Economics has also seen a growing adoption of adaptive experimentation. For example, Kasy and Sautmann (2021) and Finan and Pouzo (2022) develop new sequential experimentation strategies for use in development contexts. More recently, Chapman et al. (2024) introduced an optimal dynamic strategy for eliciting economic preferences.

Despite these growing number of applications, determining optimal decision rules in adaptive experiments remains challenging due to the interactive nature of the data generating process. For instance, there is currently no well-established notion of optimal point estimation following an adaptive experiment, nor of inference procedures that remain valid at any point during the experiment. In this article, we address these challenges by deriving a continuous-time asymptotic representation for adaptive experiments. Specifically, we model the sequential data collection process using an empirical allocation process, which, at any given time t specifies the fraction of time allocated to observing outcomes from a particular treatment. We then show that this empirical allocation process weakly converges to an allocation process in a limiting experiment, where signals consist of multiple Gaussian diffusions with unknown drifts, each corresponding to a different treatment arm. Then, by characterizing optimal decisions in the limit experiment, we can construct asymptotically optimal decision rules in the original experiment.

The limit experiment greatly simplifies the characterization of optimal decisions due to two key properties. First, its state space is considerably smaller than that of the original experiment. In the limit experiment, the sufficient statistics are the past sample paths of the signal processes. However, in many applications, these statistics can be further reduced to just the current values of the signal and allocation processes, resulting in a state-space dimension of $2K$, where K is the number of treatments. In contrast, the state space of the actual experiment encompasses all collected observations, making it substantially larger and more complex. Second, the limit experiment is far more tractable to analyze, as the properties of Gaussian diffusions in continuous time are well understood, allowing us to easily characterize optimal decisions in that setting.

To illustrate the broad applicability of our framework, we use our representation theorem to derive optimal decision rules for several fundamental aspects of adaptive experimentation. In particular, we address: (1) the construction of optimal estimators following adaptive experiments, (2) the analysis of in-sample regret, and (3) the development of e-processes for anytime-valid inference. An e-process is a nonnegative supermartingale (under the null hypothesis) that provides a principled way to track statistical evidence against the null over the course of an experiment. We introduce the first definition of an e-process for multi-treatment adaptive experiments. This enables the design of algorithm-free anytime-valid tests, i.e., tests that maintain correct size even when the sampling strategy used in the adaptive experiment is unknown.

In each of the above applications, our asymptotic framework significantly simplifies the decision problem, making it much more tractable. For example, in the estimation problem, we find that all optimal Bayes estimators share a common form that is independent of how the experiment was conducted. Furthermore, we derive explicit expressions for these estimators under Gaussian priors.

1.1. Related literature. In an important prior work, Hirano and Porter (2023) develop an asymptotic representation theorem for batched adaptive experiments, where sampling strategies are updated only a finite number of times over the course of the experiment. In this setting, they show that the policy rule and any statistic

from the finite-sample experiment can be matched to a corresponding rule and statistic in a limit experiment involving Gaussian signals from each batch.

Our asymptotic representation theorem is conceptually very different in terms of both the scope and the formulation. It applies to fully adaptive experiments and is expressed in terms of allocation processes rather than sequences of policy rules. This shift is essential, for in fully adaptive experiments, sequences of policy rules generally fail to admit weak limits. At the same time, our theory is more specialized in certain respects. It does not directly provide asymptotic representations for arbitrary statistics; instead, it characterizes the joint evolution of the score and the empirical allocation processes, which together fully determine the limit experiment. Obtaining representations for arbitrary statistics is more difficult in our setting because there is no straightforward coupling method for continuous-time processes that respects the required informational constraints (for example, ensuring that an anytime-valid test depends only on observed data).

However, for many applications such comprehensive representation theorems are not always necessary. In practice, lower bounds on losses or risk can be established more directly using our representation theorem and standard change-of-measure arguments. For instance, in the case of point estimation, we show that the frequentist risk of any sequence of estimators is asymptotically bounded below by the risk of an estimator in the limit experiment that depends only on the terminal values of the signal and allocation processes. While this result does not establish a one-to-one mapping from finite-sample estimators to their counterparts in the limit experiment, such stronger representation is not essential for deriving lower bounds or constructing asymptotically optimal estimators.

Beyond these technical contributions, the broader value of our approach is conceptual. This article provides the first general definition of adaptive experiments in continuous time through the lens of allocation processes. This formulation not only simplifies the analysis of adaptive experiments—continuous time being more tractable than discrete time—but also yields a sharp characterization of sufficient statistics. As discussed earlier, in many applications, including point estimation and anytime-valid inference, the sufficient statistics reduce to the current values of the signal and allocation processes, resulting in a finite-dimensional

state-space representation. More generally, allocation processes offer a natural and flexible framework for representing adaptive experiments, and our formulation of e-processes is indeed most naturally articulated in terms of these processes.

In terms of the style of asymptotic approximations, this article is most closely related to Le Cam’s (1979) work on stopping time representations. We extend his framework and build on his proof techniques to handle the additional complexities introduced by adaptive sampling.

The systematic study of Gaussian diffusion approximations for adaptive experiments was initiated by Fan and Glynn (2021) and Kuang and Wager (2024), who introduced diffusion asymptotics to analyze the behavior of adaptive algorithms such as Thompson Sampling. Kalvit and Zeevi (2021) extended this approach to cover UCB algorithms, and Adusumilli (2025a) further generalized the framework to characterize optimal bandit algorithms under both Bayesian and minimax regret criteria. Adusumilli (2025a) also established that likelihood ratio processes from finite-sample adaptive experiments converge uniformly over time to their diffusion counterparts. While this convergence is algorithm-agnostic and well-suited for analyzing Bayesian decision criteria, it is insufficient for tasks such as anytime-valid inference, as it lacks a representation explicitly connecting the finite-sample algorithms to a suitable counterpart in the diffusion limit. This article fills that gap by providing precisely such a representation.

Finally, this article also builds on earlier work by this author (Adusumilli, 2023) on optimal testing following adaptive experiments. The analysis of applications such as point estimation and anytime-valid inference draws on the strategies and proof techniques developed in that work.

2. ADAPTIVE EXPERIMENTS, POLICY RULES AND ALLOCATION PROCESSES

2.1. An illustrative example. We begin with a simple illustration involving two-armed bandits to motivate our theoretical analysis.

Consider a scenario in which the goal is to identify the better-performing version of a website, denoted by variants $a = 0, 1$, each with an unknown average click-through rate $\theta^{(a)}$. The observed outcome from each variant is a binary draw from

$\text{Bernoulli}(\theta^{(a)})$. To determine the optimal variant, we run a bandit algorithm for n rounds, sequentially assigning users to one of the two alternatives.

Two of the most widely used bandit algorithms in practice are Thompson Sampling (TS) and the Upper Confidence Bound (UCB) algorithm. Both use accumulated data to guide user allocation but differ in their approach.

The TS algorithm begins with a prior distribution over each $\theta^{(a)}$. Given the Bernoulli outcome model, a Beta prior is standard; for this illustration, we will take it to be $\text{Beta}(1, 1)$. At each round, the algorithm samples a draw of $\{\theta^{(a)}\}_a$ from its posterior distribution given the past observations, and then allocates the user to the variant with the highest draw.

In contrast, the UCB algorithm computes an index

$$\widehat{\text{UCB}}_j^{(a)} = \widehat{\theta}_j^{(a)} + \sqrt{\frac{2 \ln(j/n)}{N_j^{(a)}}},$$

where $\widehat{\theta}_j^{(a)}$ is the sample mean of outcomes for variant a , and $N_j^{(a)}$ is the number of users allocated to variant a prior to round j . The user is assigned to the variant with the higher index.

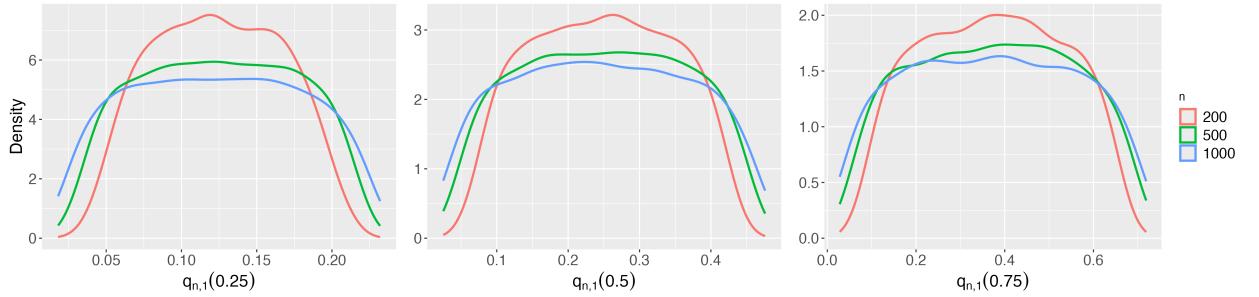
Define time $t := j/n \in [0, 1]$ as the fraction of the experiment completed. For either algorithm, let

$$q_{n,1}(t) := \frac{1}{n} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{I}\{A_j = 1\}$$

denote the number of assignments to variant 1 up to time t , normalized by n .

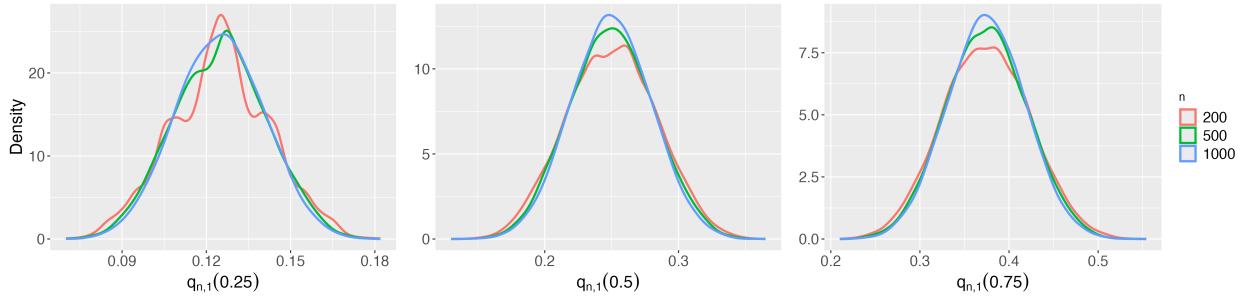
Suppose the two variants are identical, i.e., $\theta^{(0)} = \theta^{(1)}$. Figure 2.1 plots the sampling distribution of $q_{n,1}(t)$ at three time points— $t = 0.25, 0.5, 0.75$ —under TS, for various values of n . As n increases, the distributions converge, illustrating a form of asymptotic stability. The same convergence also occurs under UCB, as shown in Figure 2.2, although the sampling distributions differ quite substantially. The figures are plotted for $\theta^{(0)} = \theta^{(1)} = 0.1$, but changing these values would not make much of a difference to the plots (as long as $\theta^{(0)} = \theta^{(1)}$).

This convergence behavior of the empirical allocation process $q_{n,1}(\cdot)$ is not specific to these algorithms. A central result of this article is that under mild conditions, such asymptotic convergence holds for almost every adaptive allocation rule. To this end, we start by describing a general setup for adaptive experiments.



Note: Results from 2-armed bandit experiment with $Y^{(a)} \sim \text{Bernoulli}(\theta^{(a)})$ and $\theta^{(1)} = \theta^{(0)} = 0.1$.

FIGURE 2.1. Distribution of $q_{n,1}(t)$ under Thompson Sampling



Note: Results from 2-armed bandit experiment with $Y^{(a)} \sim \text{Bernoulli}(\theta^{(a)})$ and $\theta^{(1)} = \theta^{(0)} = 0.1$.

FIGURE 2.2. Distribution of $q_{n,1}(t)$ under UCB

2.2. General setup. Adaptive experiments involve multiple treatments, where the policy rule—i.e., the probability of allocation to each treatment—can continuously adjust over the course of the experiment.

Let K denote the number of treatments under consideration. To simplify notation and proofs, we focus on the case $K = 2$, though our results extend to any fixed K . The outcome under treatment $a \in \{0, 1\}$ follows a parametric model $\{P_{\theta^{(a)}}^{(a)}\}$, where $\theta^{(a)} \in \mathbb{R}^d$ is an unknown parameter vector. For simplicity, we assume $\theta^{(0)}$ and $\theta^{(1)}$ have the same dimension, though this is not required for our results.

Since only one outcome is observed per unit, we can assume that outcomes are independent across treatments, conditional on $(\theta^{(1)}, \theta^{(0)})$. Let $j = 1, \dots, n$ index the experimental periods. We define time t as the scaled number of periods, $t = j/n$, representing the fraction of the experiment completed.

The decision-maker (DM) employs a policy rule $\{\pi_{n,j}\}_j \equiv \{\pi_{n,\lfloor nt \rfloor}\}_t$, which prescribes the probability of assigning observation j to treatment 1 based on past information. The treatment assignment follows $A_j \sim \text{Bernoulli}(\pi_{n,j})$. For the outcomes, it is useful to conceptualize a stack of potential observations $\mathbf{y}_a :=$

$\{Y_i^{(a)}\}_{i=1}^n$ for each treatment, generated at the outset as i.i.d draws from $\{P_{\theta^{(a)}}^{(a)}\}$, but unobserved initially by the DM. This is the so-called stack-of-rewards model (Lattimore and Szepesvári, 2020, Section 4.6). Each time a treatment is sampled, it can be imagined that the DM observes the top element of the corresponding treatment stack; this element is then taken out of consideration.

2.2.1. *Empirical allocation processes.* As in the illustrative example, let

$$q_{n,a}(t) := \frac{1}{n} \sum_{j=1}^{\lfloor nt \rfloor} \mathbb{I}\{A_j = a\}$$

denote the fraction of (total) observations assigned to treatment a up to time t . We term $\{q_{n,a}(\cdot)\}_a$ the empirical allocation process. One can interpret the policy rule $\{\pi_{n,j}\}_j$ as a function mapping the stacks $(\mathbf{y}_1, \mathbf{y}_0)$ and an exogenous random variable $U \sim \text{Uniform}[0, 1]$ to the observed trajectory of $\{q_{n,a}(\cdot)\}_a$. The exogenous randomness accounts for probabilistic policy rules; specifically, we subsume the policy randomizations at all n stages of the adaptive experiment into a single U .¹ The key informational constraint is that the event $\{q_{n,1}(t) \leq \gamma_1, q_{n,0}(t) \leq \gamma_0\}$ depends only on the first $\lfloor n\gamma_1 \rfloor, \lfloor n\gamma_0 \rfloor$ observations from $(\mathbf{y}_1, \mathbf{y}_0)$ and the exogenous randomization U . Formally,

$$\{q_{n,1}(t) \leq \gamma_1, q_{n,0}(t) \leq \gamma_0\} \text{ is } \mathcal{G}_{n,\gamma_0,\gamma_1} := \mathcal{F}_{n,\gamma_1}^{(1)} \vee \mathcal{F}_{n,\gamma_0}^{(0)} \vee \sigma(U)\text{-measurable}$$

for each $t, \gamma_1, \gamma_0 \in [0, 1]$, where

$$\mathcal{F}_{n,\gamma}^{(a)} := \sigma(Y_1^{(a)}, \dots, Y_{\lfloor n\gamma \rfloor}^{(a)})$$

is the filtration (i.e., information set) generated by the first $\lfloor n\gamma \rfloor$ observations from stack \mathbf{y}_a . Thus, each policy rule $\{\pi_{n,j}\}_j$ can be associated with monotonic empirical allocation processes $\{q_{n,a}(\cdot)\}_a$ satisfying the above informational constraint.

2.2.2. *Local asymptotics.* We are interested in the behavior of various statistics under local perturbations of the form $\theta_0^{(a)} + h^{(a)}/\sqrt{n} : h^{(a)} \in \mathbb{R}^d$, where $\theta_0^{(a)}$ denotes a reference parameter. This focus is motivated by the fact that many decisions in adaptive experiments hinge on distinguishing between parameter values that are

¹Indeed, a single uniform random variable can be mapped into n independent uniform random variables.

close to one another. For example, in the website optimization setting discussed earlier, bandit algorithms are employed precisely because the differences between website variants tend to be small. Deng et al. (2013) survey industry practices and report that typical differences in click-through rates are often around 1% or less. In such cases, bandit algorithms are employed for detecting subtle effects using as little data as possible. This naturally places us in the domain of local asymptotics.

In this article, we analyze the behavior of adaptive algorithms under local asymptotics for parametric classes of families $\{P_\theta^{(a)}\}_\theta$. Let ν denote a dominating measure for $\{P_\theta^{(a)} : \theta \in \mathbb{R}^d, a \in \{0, 1\}\}$, and set $p_\theta^{(a)} := dP_\theta^{(a)}/d\nu$. We require $\{P_\theta^{(a)}\}_\theta$ to be quadratic mean differentiable (qmd):

Assumption 1. *The class $\{P_\theta^{(a)} : \theta \in \mathbb{R}^d\}$ is qmd around $\theta_0^{(a)}$ for each $a \in \{0, 1\}$, i.e., there exists a score function $\psi_a(\cdot)$ such that for each $h^{(a)} \in \mathbb{R}^d$,*

$$\int \left[\sqrt{p_{\theta_0^{(a)} + h^{(a)}}^{(a)}} - \sqrt{p_{\theta_0^{(a)}}^{(a)}} - \frac{1}{2} h^{(a)\top} \psi_a \sqrt{p_{\theta_0^{(a)}}^{(a)}} \right]^2 d\nu = o(|h^{(a)}|^2).$$

Furthermore, the information matrix $I_a := \mathbb{E}_0[\psi_a \psi_a^\top]$ is invertible for $a \in \{0, 1\}$.

In the illustrative example, the outcomes are Bernoulli, so Assumption 1 holds with $\psi_a(y) = (\theta_0^{(a)}(1 - \theta_0^{(a)}))^{-1} (y - \theta_0^{(a)})$. More broadly, this assumption is rather mild and satisfied for almost all commonly used distributions, including the Normal, Cauchy, Exponential, and Poisson distributions.

2.2.3. *Score processes.* For each $q \in [0, 1]$, define $z_{n,a}(q)$ as the partial sum process

$$z_{n,a}(q) := \frac{I_a^{-1/2}}{\sqrt{n}} \sum_{i=1}^{\lfloor nq \rfloor} \psi_a(Y_{i,j}^{(a)}).$$

Knowledge of the process, $z_{n,a}(\cdot)$, on the domain $[0, q]$ is equivalent to knowledge of the scores from the first $\lfloor nq \rfloor$ observations of the stack \mathbf{y}_a . We then define the score process for treatment a as

$$x_{n,a}(t) := z_{n,a}(q_{n,a}(t)); \quad a \in \{0, 1\}.$$

As we will show, the sample paths of this process serve as an asymptotically sufficient statistic for the adaptive experiment.

2.3. The limit experiment. The primary result of this article establishes that any adaptive experiment is asymptotically Blackwell equivalent to a limit experiment characterized by Gaussian processes.

In this limit experiment, the decision maker observes a Gaussian process signal $z_a(\cdot)$ associated with each treatment a , given by

$$z_a(q) = I_a^{1/2} h^{(a)} q + W_a(q), \quad (2.1)$$

where $\{W_a(\cdot)\}_a$ are independent d -dimensional Wiener processes, and the drifts $\{h^{(a)}\}_a$ are unknown. Intuitively, $z_a(\cdot)$ serves as the limiting counterpart of $z_{n,a}(\cdot)$ in the original experiment, while the index q represents the amount of ‘attention’ devoted to that particular treatment.

Define the natural filtration generated by the Gaussian process $z_a(\cdot)$ up to a given attention-value γ as $\mathcal{F}_\gamma^{(a)} := \sigma\{z_a(s) : s \leq \gamma\}$. We can and do take $\mathcal{F}_\gamma^{(a)}$ to be right-continuous, i.e., $\mathcal{F}_\gamma^{(a)} \equiv \bigcap_{\epsilon \downarrow 0} \mathcal{F}_{\gamma+\epsilon}^{(a)}$. Similar to the stack of rewards in the original experiment, the entire process $z_a(\cdot)$ is not immediately observed. Instead, at time t , the DM observes the sample paths of $z_1(\cdot)$ and $z_0(\cdot)$ over the intervals $[0, q_1(t)]$ and $[0, q_0(t)]$, respectively, where $q_a(t)$ represents the amount of attention devoted to treatment a up to time t . The quantities $\{q_a(t)\}_a$, termed allocation processes, are continuous-time analogues of $\{q_{n,a}(t)\}_a$, and formally defined as follows:

Definition 1. Let $(z_1(\cdot), z_0(\cdot), U)$ represent a collection of independent stochastic processes and random variables defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\{z_a(\cdot)\}_a$ are defined as in (2.1) and U is an exogenous Uniform $[0, 1]$ random variable. A collection of non-negative stochastic processes, $\{q_a(\cdot)\}_a$, indexed by $t \in [0, 1]$, is termed an **allocation process** if:

- (i) With probability 1, $q_1(t) + q_0(t) = t \forall t$;
- (ii) Each $q_a(\cdot)$ is almost surely non-decreasing; and
- (iii) For any $\gamma_1, \gamma_0, t \in [0, 1]$ such that $\gamma_1 + \gamma_0 \geq t$, the event $\{q_1(t) \leq \gamma_1, q_0(t) \leq \gamma_0\}$ is measurable with respect to $\mathcal{G}_{\gamma_1, \gamma_0}$, the augmented version of the filtration $\mathcal{F}_{\gamma_1}^{(1)} \vee \mathcal{F}_{\gamma_0}^{(0)} \vee \sigma(U)$.²

²The augmented version of $\mathcal{F}_{\gamma_1}^{(1)} \vee \mathcal{F}_{\gamma_0}^{(0)} \vee \sigma(U)$ is the smallest filtration containing $\mathcal{F}_{\gamma_1}^{(1)} \vee \mathcal{F}_{\gamma_0}^{(0)} \vee \sigma(U)$ that includes every null set of $(\Omega, \mathcal{F}, \mathbb{P})$, i.e., every $A \in \mathcal{F}$ such that $\mathbb{P}(A) = 0$.

The first two conditions on $q_a(\cdot)$ are straightforward. The third condition ensures that, aside from the exogenous randomization U , whether or not $q_1(t) \leq \gamma_1$ and $q_0(t) \leq \gamma_0$ hold can be determined based only on the sample paths, $\{z_a(s) : s \leq \gamma_a\}$, of the cumulative past outcomes. It is the continuous time counterpart of the information constraint on empirical allocation processes, described in Section 2.2.1. The condition is analogous to the usual definition of a stopping time, but extended to a multi-dimensional setting.

Although the informational constraint is only explicitly required for $\gamma_1 + \gamma_0 \geq t$, the requirement $q_1(t) + q_0(t) = t$ almost surely implies that the event $\{q_1(t) \leq \gamma_1, q_0(t) \leq \gamma_0\}$ is a measure-zero set when $\gamma_1 + \gamma_0 < t$. Consequently, because $\mathcal{G}_{\gamma_1, \gamma_0}$ is an augmented filtration, this event is measurable with respect to $\mathcal{G}_{\gamma_1, \gamma_0}$ for any $\gamma_1 + \gamma_0 < t$ as well.

The quantities $\{z_a(\cdot), q_a(\cdot), U\}_a$ characterize the limit adaptive experiment. Importantly, the inputs to the processes $z_a(\cdot), q_a(\cdot)$ are different: $\{q_a(\cdot)\}_a$ are indexed by time, while $\{z_a(\cdot)\}_a$ are indexed by the attention devoted to each treatment.

2.3.1. Sufficient statistics. For each $a \in \{0, 1\}$, define $x_a(t) := z_a(q_a(t))$ as the limit counterpart of $x_{n,a}(t)$. It is straightforward to verify that the sample paths of $\{x_a(\cdot)\}_a$ constitute sufficient statistics for the limit experiment up to time t .

Let $\mathcal{I}_t := \mathcal{G}_{q_1(t), q_0(t)}$ denote the information accrued from the experiment until time t .³ An important property of $x_1(\cdot), x_0(\cdot)$ is that they are \mathcal{I}_t -martingales when $\mathbf{h} := (h^{(1)}, h^{(0)}) = (0, 0)$. Furthermore, $q_a(t)$ is the quadratic variation of $x_a(t)$, in that it captures the accumulated variability of $x_a(t)$.

Lemma 1. *Under $\mathbf{h} = (0, 0)$, the processes $x_1(\cdot), x_0(\cdot)$ are \mathcal{I}_t -martingales with quadratic variations $q_1(t), q_0(t)$.*

2.3.2. Allocation processes vs policy rules. Theorem 1 in Section 3 states that we can take $q_a(\cdot)$ to be almost surely Lipschitz continuous with a Lipschitz constant of 1. By the fundamental theorem of Lebesgue integral calculus, almost every sample path of $q_a(\cdot)$ is then differentiable almost everywhere, with a Lebesgue integrable derivative $\pi_a(t) := \frac{dq_a(t)}{dt}$. Consequently, $q_a(t) = \int_0^t \pi_a(s) ds$. Moreover,

³Formally, $\mathcal{G}_{q_1(t), q_0(t)} \equiv [A \in \mathcal{F} : A \cap \{q_1(t) \leq \gamma_1, q_0(t) \leq \gamma_0\} \in \mathcal{G}_{\gamma_1, \gamma_0} \forall t, \gamma_1, \gamma_0 \in [0, 1]]$.

Definition 1 and the right continuity of \mathcal{G}_{q_1, q_0} implies that $\pi_a(t)$ is \mathcal{I}_t -measurable coordinate-wise, for each t .

However, $\pi(\cdot) \equiv \pi_1(\cdot)$ cannot, in general, be interpreted as a valid policy rule, since it need not be measurable as a random process indexed by t . Indeed, in many continuous-time optimal control problems, the optimal policy does not belong to a separable space, and is therefore not measurable.⁴ For this reason, it is also generally impossible to define weak convergence for a sequence of policies $\pi_{n, \lfloor nt \rfloor}$, as such sequences are typically not asymptotically equicontinuous.

These observations indicate that the policy rule is not a particularly suitable object for characterizing sequential strategies in the limit experiment. In contrast, as will be established in Theorem 1 below, any sequence of empirical allocation processes converge weakly to an allocation process in the limit experiment. In this sense, the allocation process serves as the more fundamental representation of sequential decision-making.

3. ASYMPTOTIC EQUIVALENCE OF EXPERIMENTS

We now establish the asymptotic equivalence between the original sequence of adaptive experiments and the limit experiment. This equivalence follows from two key results. The first, previously demonstrated in Adusumilli (2025a), states that the likelihood ratio processes in the original experiment converge uniformly to their counterparts in the limit experiment. We restate this result here for completeness. The second result, which is novel to this article, asserts that any sequence of score and allocation processes, $\{x_{n,a}(\cdot), q_{n,a}(\cdot)\}_a$, admits a corresponding representation in the limit experiment.

3.1. Convergence of likelihood ratio processes. Let $\mathbb{P}_{n,h}$ denote the induced probability over the stacked rewards and the exogenous randomization, i.e., over $(\mathbf{y}^{(1)}, \mathbf{y}^{(0)}, U)$ when $Y^{(a)} \sim P_{\theta_0 + h^{(a)}/\sqrt{n}}^{(a)}$. For each a , denote $\mathbf{y}_{\lfloor nq \rfloor}^{(a)} := \{Y_i^{(a)}\}_{i=1}^{\lfloor nq \rfloor}$. Suppose that we observe U and $\lfloor n\gamma_1 \rfloor, \lfloor n\gamma_0 \rfloor$ units from each treatment, i.e., we

⁴The measurability of sample paths—often referred to as strong or Bochner measurability—differs from the weaker notion of coordinate-wise measurability. The Pettis measurability theorem states that a stochastic process is Bochner measurable if and only if its sample paths lie in a separable subspace with probability one. See Adusumilli (2023, Section 5.1.1) for further discussion.

observe U and $\mathbf{y}_{\gamma_1, \gamma_0} := (\mathbf{y}_{\lfloor n\gamma_1 \rfloor}^{(1)}, \mathbf{y}_{\lfloor n\gamma_0 \rfloor}^{(0)})$. Given this set of observations, the log-likelihood ratio process under the local alternative $\mathbf{h} := (h^{(1)}, h^{(0)})$ is:

$$\begin{aligned}\hat{\varphi}(\mathbf{h}; \gamma_1, \gamma_0) &= \ln \frac{dP_{\theta_0^{(1)} + h^{(1)}/\sqrt{n}}^{(1)}}{dP_{\theta_0^{(1)}}^{(1)}} (\mathbf{y}_{\lfloor n\gamma_1 \rfloor}^{(1)}) + \ln \frac{dP_{\theta_0^{(0)} + h^{(0)}/\sqrt{n}}^{(0)}}{dP_{\theta_0^{(0)}}^{(0)}} (\mathbf{y}_{\lfloor n\gamma_0 \rfloor}^{(0)}) \\ &:= \hat{\varphi}^{(1)}(\mathbf{h}; \gamma_1) + \hat{\varphi}^{(0)}(\mathbf{h}; \gamma_0),\end{aligned}$$

where, for any $a \in \{0, 1\}$ and $\gamma \in [0, 1]$,

$$\ln \frac{dP_{\theta_0^{(a)} + h^{(a)}/\sqrt{n}}^{(a)}}{dP_{\theta_0^{(a)}}^{(a)}} (\mathbf{y}_{\lfloor n\gamma \rfloor}^{(a)}) := \sum_{i=1}^{\lfloor n\gamma \rfloor} \ln \frac{dP_{\theta_0^{(a)} + h^{(a)}/\sqrt{n}}^{(a)}}{dP_{\theta_0^{(a)}}^{(a)}} (Y_i^{(a)}).$$

In Adusumilli (2025a), this author showed that under Assumption 1,

$$\hat{\varphi}^{(a)}(\mathbf{h}; \gamma) = h^{(a)\top} I_a^{1/2} z_{n,a}(\gamma) - \frac{\gamma}{2} h^{(a)\top} I_a h^{(a)} + o_{\mathbb{P}_{n,0}}(1) \text{ uniformly over } \gamma \in [0, 1]. \quad (3.1)$$

Analogously, in the limit experiment, the relevant probability measure is $\mathbb{P}_h := \mathbb{P}_{h^{(1)}}^{(1)} \otimes \mathbb{P}_{h^{(0)}}^{(0)} \otimes P_U$, where $\mathbb{P}_{h^{(a)}}^{(a)}$ is the induced probability over the sample paths of $\{z_a(s); 0 \leq s \leq 1\}$ when the local parameter is $h^{(a)}$, and P_U is the probability measure induced by $U \sim \text{Uniform}[0, 1]$. Also, for some fixed $\{\gamma_a\}_a$, let

$$\varphi^{(a)}(h^{(a)}; \gamma_a) = \mathbb{E}_{\mathbb{P}_0^{(a)}} \left[\ln \frac{d\mathbb{P}_{h^{(a)}}^{(a)}}{d\mathbb{P}_0^{(a)}} \middle| \bar{\mathcal{F}}_{\gamma_a}^{(a)} \vee \sigma(U) \right]$$

denote the log-likelihood ratio under the local alternative $h^{(a)}$ given the sample path $\{z_a(s); s \leq \gamma_a\}$ and U . Similarly, $\varphi(\mathbf{h}; \gamma_1, \gamma_0)$ denotes the likelihood ratio under $\mathbf{h} = (h^{(1)}, h^{(0)})$ given the sample paths $\{z_a(s); s \leq \gamma_a\}_a$ and U . Since $z_1(\cdot), z_0(\cdot)$ are Wiener processes under \mathbb{P}_0 , the Girsanov theorem implies

$$\varphi^{(a)}(h^{(a)}; \gamma_a) = h^{(a)\top} I_a^{1/2} z_a(\gamma_a) - \frac{\gamma_a}{2} h^{(a)\top} I_a h^{(a)}. \quad (3.2)$$

Furthermore, as the Wiener processes are independent, $\varphi(\mathbf{h}; \gamma_1, \gamma_0) = \varphi^{(1)}(h^{(1)}; \gamma_1) + \varphi^{(0)}(h^{(0)}; \gamma_0)$.

Equations (3.1), (3.2) imply $\hat{\varphi}^{(a)}(h^{(a)}; \cdot) \xrightarrow[\mathbb{P}_{n,0}]{} \varphi^{(a)}(h^{(a)}; \cdot)$ for each $a, h^{(a)}$ and therefore,

$$\hat{\varphi}(\mathbf{h}; \cdot, \cdot) \xrightarrow[\mathbb{P}_{n,0}]{} \varphi(\mathbf{h}; \cdot, \cdot) \text{ for each } \mathbf{h}. \quad (3.3)$$

In other words, the likelihood ratio processes under the original experiment converge uniformly to those under the limit experiment for each \mathbf{h} . Following Blackwell (1953) and Le Cam (1986), this implies that the two experiments are equivalent, in that the posteriors converge uniformly over all possible values of γ_1, γ_0 .

By itself, (3.3) does not make any reference to an allocation process. To show that a risk function in the original experiment admits a corresponding representation or lower bound in the limit experiment, we need to go further and match the allocation processes as well. In particular, we need to show that for any set of (asymptotically) sufficient statistics, $\{x_{n,a}(\cdot), q_{n,a}(\cdot)\}_a$, there exists a corresponding $\{x_a(\cdot), q_a(\cdot)\}_a$ in the limit experiment such that, when $\mathbf{h} = (0, 0)$, the distribution of the former converges to that of the latter as $n \rightarrow \infty$. This is the key result that we prove next.

3.2. The main result.

Theorem 1. *Suppose Assumption 1 holds. Let $\{x_{n,a}(\cdot), q_{n,a}(\cdot)\}_a$ be any sequence of score and allocation processes induced by a sequence of policies $\{\pi_{n,j}\}_j$ in the actual experiment. Then, there exists a further subsequence, $\{n_k\}_{k=1}^\infty$, and a random collection $\{z_a(\cdot), q_a(\cdot), U\}_a$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that:*

- (i) $\{z_a(\cdot)\}_a$ are independent standard d -dimensional Wiener processes and $U \sim \text{Uniform}[0, 1]$ is independent of $\{z_a(\cdot)\}_a$;
- (ii) $\{q_a(\cdot)\}_a$ is an allocation process in the sense of Definition 1;
- (iii) $\{x_{n_k,a}(\cdot), q_{n_k,a}(\cdot)\}_a \xrightarrow[\mathbb{P}_{n,0}]{} \{x_a(\cdot), q_a(\cdot)\}_a$, where $x_a(t) := z_a(q_a(t))$; and
- (iv) $\{q_a(\cdot)\}_a$ is almost surely Lipschitz continuous, with a Lipschitz constant of 1.

Theorem 1 establishes that the distribution of $\{x_{n_k,a}(\cdot), q_{n_k,a}(\cdot)\}_a$ in the original experiment can be matched with that of $\{x_a(\cdot), q_a(\cdot)\}_a$ in the limit experiment, where $\{q_a(\cdot)\}_a$ is a suitably defined allocation process. Although these statistics are path-valued processes, the convergence is uniform over time. In concert with (3.3), Theorem 1 enables us to derive lower bounds on losses or statistical risk in various applications by employing change of measure arguments.

While the proof of Theorem 1 is somewhat involved, the underlying intuition is relatively straightforward. The processes $\{z_{n,a}(\cdot)\}_a$ are asymptotically tight (being standard partial sum processes), and likewise, the processes $\{q_{n,a}(\cdot)\}_a$ are

also tight since, by definition,

$$\sup_t |q_{n,a}(t + \delta) - q_{n,a}(t)| \leq \delta + n^{-1}, \quad \forall \delta > 0.$$

This ensures that $\{z_{n,a}(\cdot), q_{n,a}(\cdot)\}_a$ converges to some weak limit $\{z_a(\cdot), q_a(\cdot)\}_a$. Moreover, the measurability of the events $\{q_{n,1}(t) \leq \gamma_1, q_{n,0}(t) \leq \gamma_0\}$ with respect to $\mathcal{G}_{n,\gamma_0,\gamma_1} := \mathcal{F}_{n,\gamma_1}^{(1)} \vee \mathcal{F}_{n,\gamma_0}^{(0)} \vee \sigma(U)$ suggests that $q_a(\cdot)$ can be constructed to inherit the appropriate adaptedness properties required by Definition 1. The construction requires some care and utilizes some results from the theory of stable convergence (Häusler and Luschgy, 2015); it is perhaps the most intricate part of the proof. Setting $x_{n,a}(\cdot) := z_{n,a}(q_{n,a}(\cdot))$ and $x_a(\cdot) := z_a(q_a(\cdot))$ then gives the desired result.

Although Theorem 1 is stated for sub-sequences, most applications require weak convergence of the full sequence $\{x_{n,a}(\cdot), q_{n,a}(\cdot)\}_a$. This motivates:

Assumption 2. *The sequence of policy rules $\{\pi_{n,j}\}_j$ is such that $\{x_{n,a}(\cdot), q_{n,a}(\cdot)\}_a$ has a weak limit under $\mathbb{P}_{n,0}$.*

Theorem 1 already ensures $\{x_{n,a}(\cdot), q_{n,a}(\cdot)\}_a$ is tight under $\mathbb{P}_{n,0}$. Assumption 2 strengthens this to weak convergence. The assumption is needed to rule out pathological sequences of policy rules, e.g., sequences where the policy rules differ for even and odd n . It is therefore rather mild: if it does not hold, one should extract convergent subsequences and treat each as arising from a distinct protocol.

3.3. Behavior under local alternatives. Theorem 1 describes the behavior of $\{x_{n_k,a}(\cdot), q_{n_k,a}(\cdot)\}_a$ under the reference distribution $\mathbb{P}_{n,0}$. Under local alternatives of the form $\mathbb{P}_{n,h}$, the partial sum processes, $z_{n,a}(\cdot)$, acquire an asymptotic drift, converging weakly to $z_a(\cdot) \sim I_a^{1/2} h_a \cdot + W_a(\cdot)$.

Given that $\{x_a(\cdot), q_a(\cdot)\}_a$ is adapted to the filtrations generated by $\{z_a(\cdot)\}_a$ and the exogenous randomization $U \sim \text{Uniform}[0, 1]$, it follows that $\{x_{n,a}(\cdot), q_{n,a}(\cdot)\}_a$ should converge weakly to $\{x_a(\cdot), q_a(\cdot)\}_a$, where the only difference is that the underlying processes $\{z_a(\cdot)\}_a$ now exhibit a linear drift. This is formalized in the following corollary.

Corollary 1. *Suppose Assumptions 1 and 2 hold. Let $\{x_{n,a}(\cdot), q_{n,a}(\cdot)\}_a$ be any sequence of score and allocation processes induced by a sequence of policies $\{\pi_{n,j}\}_j$*

in the actual experiment. Then, there exists a random collection $\{z_a(\cdot), q_a(\cdot), U\}_a$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that:

- (i) $z_a(\cdot) \sim I_a^{1/2} h^{(a)} \cdot + W_a(\cdot)$ are independent Gaussian processes and $U \sim \text{Uniform}[0, 1]$ is independent of $\{z_a(\cdot)\}_a$;
- (ii) $\{q_a(\cdot)\}_a$ —which is invariant across \mathbf{h} as a function of $(U, \{z_a(s) : 0 \leq s \leq 1\}_a)$ —is an allocation process in the sense of Definition 1; and
- (iii) $\{x_{n,a}(\cdot), q_{n,a}(\cdot)\}_a \xrightarrow[\mathbb{P}_{n,\mathbf{h}}]{d} \{x_a(\cdot), q_a(\cdot)\}_a$, where $x_a(t) := z_a(q_a(t))$.

Corollary 1 is established in Appendix A using Theorem 1 and standard change-of-measure arguments analogous to Le Cam’s third lemma.

Theorem 1 and Corollary 1 are existence results: they establish that $\{q_{n,a}(\cdot)\}_a$ converge weakly to an allocation process $\{q_a(\cdot)\}_a$ in the limit experiment. While these results do not characterize the explicit form of $\{q_a(\cdot)\}_a$, this is generally not a limitation in practice. As the applications below illustrate, it is often possible to determine the form of optimal decisions in the limit experiment without knowing the specific structure of $\{q_a(\cdot)\}_a$. Theorem 1, combined with change-of-measure arguments, then allows us to transfer these decisions back to the finite-sample setting and show that they are asymptotically optimal.

4. APPLICATION 1: POINT ESTIMATION

In this section, we illustrate how Theorem 1 can be used to analyze estimation problems following an adaptive experiment.

Suppose that, upon completion of the experiment, we aim to estimate the unknown parameter vector $\boldsymbol{\theta} := (\theta^{(1)}, \theta^{(0)})$. Let T_n denote a proposed estimator based on the entire data collected before the terminal time $t = 1$. By construction, T_n must be $\mathcal{I}_{n,1} := \mathcal{G}_{n,q_{n,1}(1),q_{n,0}(1)}$ measurable.

Let $l(\cdot)$ be a non-negative convex loss function. Following the setup introduced earlier in this paper, we fix a reference parameter $\boldsymbol{\theta}_0$ and evaluate estimator loss under local alternatives of the form $\boldsymbol{\theta}_0 + \mathbf{h}/\sqrt{n}$, where $\mathbf{h} \in \mathbb{R}^d$. Unlike classical settings, however, the choice of $\boldsymbol{\theta}_0$ is subject to important restrictions. In many adaptive experiments, only certain reference points yield non-degenerate diffusion asymptotics. Let Θ_0 denote the equivalence class of such admissible reference parameters. For example, in two-armed bandit experiments, this class consists

of all parameter vectors satisfying $\theta^{(1)} = \theta^{(0)}$; otherwise, the resulting allocation processes become asymptotically degenerate, collapsing to either 0 or 1. More generally, we define Θ_0 to be the set of all parameter values $\boldsymbol{\theta}$ for which the weak limit of $q_{n,a}(\cdot)$ is not trivial (i.e., not identically 0) for any arm a . The reference parameter $\boldsymbol{\theta}_0$ must therefore lie in this set; otherwise some components of $\boldsymbol{\theta}$ will not be consistently estimable.⁵

Given such a choice of $\boldsymbol{\theta}_0$, the frequentist risk of T_n , evaluated at the local parameter \mathbf{h} , is defined as

$$R_n(T_n, \mathbf{h}) = \mathbb{E}_{n,\mathbf{h}} \left[l \left(\sqrt{n}(T_n - \boldsymbol{\theta}(\mathbf{h})) \right) \right],$$

where $\mathbb{E}_{n,\mathbf{h}}[\cdot]$ is the expectation under $\mathbb{P}_{n,\mathbf{h}}$, and $\boldsymbol{\theta}(\mathbf{h}) := \boldsymbol{\theta}_0 + \mathbf{h}/\sqrt{n}$.

The estimation problem in the limit experiment is defined analogously. Given access to the sample paths of $\{x_a(\cdot), q_a(\cdot)\}_a$ over $t \in [0, 1]$, we seek an estimate of the local parameter \mathbf{h} . Let T denote a candidate estimator, required to be $\mathcal{I}_1 \equiv \mathcal{G}_{q_1(1), q_0(1)}$ measurable. The frequentist risk of this estimator is

$$R(T, \mathbf{h}) := \mathbb{E}_{\mathbf{h}} [l(T - \mathbf{h})].$$

We term a sequence of estimators, $\{T_n\}_n$, *tight* at $\boldsymbol{\theta}_0$ if $\sqrt{n}(T_n - \boldsymbol{\theta}_0)$ is asymptotically tight, i.e., bounded in probability, under $\mathbb{P}_{n,0}$. Tightness at $\boldsymbol{\theta}_0$ is a substantial relaxation of the usual notion of regularity—which requires the limit distribution of $\sqrt{n}(T_n - \boldsymbol{\theta}(\mathbf{h}))$ under $\mathbb{P}_{n,\mathbf{h}}$ to be the same for all \mathbf{h} .

The following theorem states that the asymptotic performance of any tight sequence of estimators $\{T_n\}_n$ is lower bounded, along subsequences, by the performance of some estimator T in the limit experiment, and that this limit estimator depends only on the terminal values $\{x_a(1), q_a(1)\}_a$.

Theorem 2. *Under Assumptions 1-2, for any tight sequence of estimators, $\{T_n\}_n$, there exists a further sub-sequence, $\{T_{n_k}\}_k$, and an estimator T in the limit experiment depending only on $\{x_a(1), q_a(1)\}_a$ such that $\liminf_{k \rightarrow \infty} R_{n_k}(T_{n_k}, \mathbf{h}) \geq R(T, \mathbf{h})$ for each \mathbf{h} .*

⁵The set Θ_0 can be enlarged, however, if we are only interested in estimating some sub-components of $\boldsymbol{\theta}$.

The surprising implication of Theorem 2 is that knowledge of just the terminal values of $\{x_a(\cdot), q_a(\cdot)\}_a$ is sufficient to characterize optimal estimators. The paths taken by these processes are not directly informative for estimation.

4.1. Bayes risk. Let $m_{\theta}(\cdot)$ denote a given prior over θ . This induces a local prior, $m(\cdot)$, over \mathbf{h} through the transformation $\mathbf{h} = \sqrt{n}(\theta - \theta_0)$. We consider an asymptotic regime wherein $m(\cdot)$ is assumed to be independent of n . The influence of the prior thus remains asymptotically non-negligible and the Bernstein-von Mises theorem does not apply. As discussed in Adusumilli (2025a), local priors offer a better approximation to finite-sample behavior because their influence does not diminish with sample size.

Theorem 2 implies a lower bound on the Bayes risk corresponding to $m(\cdot)$:

Corollary 2. *Under Assumptions 1-2, for any tight sequence of estimators, $\{T_n\}_n$, there exists a further subsequence, $\{T_{n_k}\}_k$, and an estimator T in the limit experiment depending only on $\{x_a(1), q_a(1)\}_a$ such that the Bayes risk, $\int R_{n_k}(T_{n_k}, \mathbf{h})dm(\mathbf{h})$, of $\{T_{n_k}\}_k$ is asymptotically lower bounded by the Bayes risk, $\int R(T, \mathbf{h})dm(\mathbf{h})$, of T in the limit experiment.*

Let T^* denote the optimal Bayes estimator in the limit experiment. Then, Corollary 2 implies that $R^*(m) := R(T^*, m)$ is an asymptotic lower bound on the Bayes risk of any tight sequence of estimators T_n , i.e.,

$$\liminf_{n \rightarrow \infty} R_n(T_n, m) \geq R^*(m) \quad \forall T_n.$$

This lower bound does not require the use of subsequences.

By the likelihood principle, the optimal Bayes estimator in the limit experiment is algorithm independent and depends only on $\{x_a(1), q_a(1)\}_a$. For example, consider a prior $m_0(\cdot)$ on $\mathbf{h} \equiv (h^{(1)}, h^{(0)}) \in \mathbb{R}^2$ with independent Gaussian components: $\mathcal{N}(\mu_0^{(1)}, \nu_{(1)}^2) \times \mathcal{N}(\mu_0^{(0)}, \nu_{(0)}^2)$. Then, by standard results in stochastic filtering, the posterior distribution of $h^{(a)}$ at the end of the experiment is

$$h^{(a)} | q_a(1), x_a(1) \sim \mathcal{N} \left(\frac{I_a^{1/2} x_a(1) + \nu_{(a)}^{-2} \mu_0^{(a)}}{I_a q_a(1) + \nu_{(a)}^{-2}}, \frac{1}{I_a q_a(1) + \nu_{(a)}^{-2}} \right),$$

for each a . The optimal Bayes estimator of $h^{(a)}$ in the limit experiment, under the squared error loss, $l(\delta) = \delta^2$, is therefore

$$T^* = \frac{I_a^{1/2}x_a(1) + \nu_{(a)}^{-2}\mu_0^{(a)}}{I_aq_a(1) + \nu_{(a)}^{-2}}.$$

Notably, this estimator is invariant to the choice of the sampling algorithm. As $\nu^{(a)} \rightarrow \infty$, we get the MLE estimator $T_{\text{mle}} = I_a^{-1/2}x_a(1)/q_a(1)$.

Corollary 2 implies that the set of estimators depending only on $\{x_a(1), q_a(1)\}_a$ constitute a complete class in the limit experiment. Furthermore, we can lower bound the Bayes risk of any tight sequence of estimators using the Bayes risk of estimators in the limit experiment. These results are useful because determining the optimal estimator is a lot easier in the limit experiment. As seen above, Gaussian priors are particularly straightforward to analyze due to conjugacy. For general priors, computing the posterior is more involved, but one can employ approximate methods such as MCMC.

4.2. Attaining the bound. Given an optimal Bayes estimator, $T^*(\{x_a(1), q_a(1)\}_a)$, in the limit experiment, we can construct a finite sample version,

$$T_n^* := \boldsymbol{\theta}_0 + n^{-1/2}T^*(\{x_{n,a}(1), q_{n,a}(1)\}_a),$$

by replacing $x_a(\cdot), q_a(\cdot)$ with the sample counterparts $x_{n,a}(\cdot), q_{n,a}(\cdot)$. Since T^* is an estimator of \mathbf{h} , the transformation above converts it into an estimator of $\boldsymbol{\theta}$.

In practice, because $x_{n,a}(\cdot)$ depends on the information matrix, I_a , one would need to replace it with a consistent estimate. This can be supplied by the standard variance estimator, which remains consistent under general conditions, even if only at slower-than $n^{-1/2}$ rates. The construction also requires knowledge of the reference parameter $\boldsymbol{\theta}_0$. We suggest choosing this as the element from the equivalence class, $\boldsymbol{\Theta}_0$, that is closest to the prior median under $m_{\boldsymbol{\theta}}(\cdot)$.

Suppose that $T^*(\cdot)$ satisfies the conditions for a continuous mapping theorem. Together with (3.1) and Theorem 1, this implies

$$\begin{pmatrix} \sqrt{n}(T_n^* - \boldsymbol{\theta}(\mathbf{h})) \\ \hat{\varphi}(\mathbf{h}; q_{n,1}(1), q_{n,0}(1)) \end{pmatrix} \xrightarrow{P_{nT,0}} \begin{pmatrix} T^* - h \\ \sum_a \left\{ h^{(a)\top} I_a^{1/2} x_a(1) - \frac{q_a(1)}{2} h^{(a)\top} I_a h^{(a)} \right\} \end{pmatrix},$$

for any \mathbf{h} . Then, a similar argument as in the proof of Theorem 2 shows that the frequentist risk of T_n^* converges to that of T^* in the limit experiment, as long as the loss function is bounded. But T^* is the optimal Bayes estimator in the limit experiment, so the above implies that T_n^* is asymptotically Bayes optimal as well, in the sense that its Bayes risk is arbitrarily close to $R^*(m)$ as $n \rightarrow \infty$.

4.3. Minimax risk. Minimax risk is defined as $\inf_{T_n} \sup_{m(\cdot)} \int R_n(T_n, \mathbf{h}) dm(\mathbf{h})$, where T_n ranges over all tight sequences of estimators, and $m(\cdot)$ ranges over all possible priors. The estimator T_n^* that solves this problem is referred to as the minimax estimator. This estimator can also be interpreted as the equilibrium outcome of a zero-sum game between a decision-maker and nature: nature selects a prior $m(\cdot)$ to maximize the Bayes risk, while the decision-maker selects an estimator to minimize it.

In contrast to classical experiments, the minimax risk in adaptive experiments is often infinite. To see why, consider the two-armed bandit experiment from the illustrative example, and suppose the objective is to estimate $h^{(1)}$ in the limit experiment. Nature can make the problem arbitrarily hard by choosing a flat prior over $h^{(1)}$ and taking $h^{(0)} \rightarrow \infty$. In this case, because $h^{(0)}/h^{(1)} \rightarrow \infty$ with probability one, the bandit algorithm devotes negligible attention to arm 1, effectively yielding no data from which to estimate $h^{(1)}$. This leads to an infinite risk.

4.4. Illustrative example (contd.) Continuing with the illustrative example from Section 2.1, suppose we aim to estimate the parameter $\theta^{(1)}$ after conducting the experiment using a bandit algorithm (we show results under both UCB and Thompson Sampling). We assume independent and identical Gaussian priors over $\theta^{(0)}$ and $\theta^{(1)}$, given by $\mathcal{N}(\bar{\theta}, \bar{\sigma}^2)$.

To apply our asymptotic framework, we reformulate the problem as a local estimation problem. Specifically, we treat $\boldsymbol{\theta}_0 = (\bar{\theta}, \bar{\theta})$ —the vector of prior medians—as the reference value and define the local parameter $h^{(a)} := \sqrt{n}(\theta^{(a)} - \bar{\theta})$. This transformation induces a prior over $h^{(a)}$ of the form $\mathcal{N}(0, \nu^2)$, where $\nu^2 := n\bar{\sigma}^2$. As detailed in Section 4.1, ν^2 is held fixed in our asymptotic regime even as n increases, implying that the prior over $\theta^{(a)}$ increasingly concentrates around $\bar{\theta}$.

In the Bernoulli case, the score function is $\psi(y) = [\bar{\theta}(1 - \bar{\theta})]^{-1} (y - \bar{\theta})$, and the Fisher information is $I = [\bar{\theta}(1 - \bar{\theta})]^{-1}$. Then,

$$x_{n,a}(t) := \frac{1}{\sqrt{n\bar{\theta}(1 - \bar{\theta})}} \sum_{j=1}^{\lfloor nq_{n,a}(t) \rfloor} (Y_j^{(a)} - \bar{\theta}).$$

As shown in Section 4.2, the asymptotically optimal Bayes estimator of $h^{(1)}$ is

$$\hat{h}^{(1)} = (I q_{n,1}(1) + \nu^{-2})^{-1} I^{1/2} x_{n,1}(1).$$

This leads to the corresponding estimator for $\theta^{(1)}$:

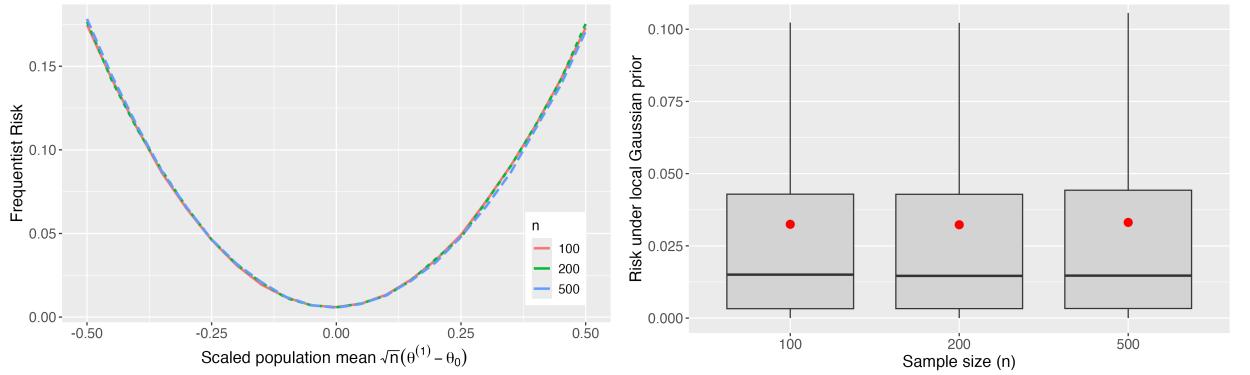
$$\hat{\theta}^{(1)} = \frac{I^{-1}\nu^{-2}}{q_{n,1}(1) + I^{-1}\nu^{-2}} \bar{\theta} + \frac{1}{q_{n,1}(1) + I^{-1}\nu^{-2}} \left(\frac{1}{n} \sum_{j=1}^{\lfloor nq_{n,1}(1) \rfloor} Y_j^{(1)} \right).$$

Clearly, $\hat{\theta}^{(1)}$ has the same form as the usual shrinkage estimator of the population mean under a Gaussian prior. As $\nu \rightarrow \infty$, $\hat{\theta}^{(1)}$ becomes the MLE estimator $\frac{1}{nq_{n,1}(1)} \sum_{j=1}^{\lfloor nq_{n,1}(1) \rfloor} Y_j^{(1)}$. Both estimators are independent of the adaptive sampling algorithm used.

To evaluate the finite-sample performance of $\hat{\theta}^{(1)}$, we conduct simulations based on the illustrative example from Section 2.1, fixing $\theta^{(0)} = \bar{\theta}$ and setting $\theta^{(1)} = \bar{\theta} + h^{(1)}/\sqrt{n}$ for values of $h^{(1)}$ ranging from -0.5 to 0.5 . We also take $\bar{\theta} = 0.1$ and set the prior standard deviation over $h^{(1)}$ to be $\nu = 0.2$. This implies that the 95% prior credible interval for $\theta^{(1)}$ is approximately $[0.04, 0.14]$ when $n = 100$.

Panel A of Figure 4.1 displays the corresponding frequentist risk profiles of the estimator for different values of n , when the data is obtained through UCB. Notably, the risk profiles are nearly identical across sample sizes, even when n is as small as 100, highlighting the robustness of the estimator's performance in small samples. Panel B of the same figure plots the resulting Bayes risk under the local prior $(h^{(1)}, h^{(0)}) \sim \text{i.i.d } \mathcal{N}(0, \nu^2)$. The distributions of risk are again almost identical across n .

Figure 4.2 displays equivalent results when the data is obtained through Thompson Sampling. Surprisingly, the estimator attains very similar values of Bayes risk under both algorithms.

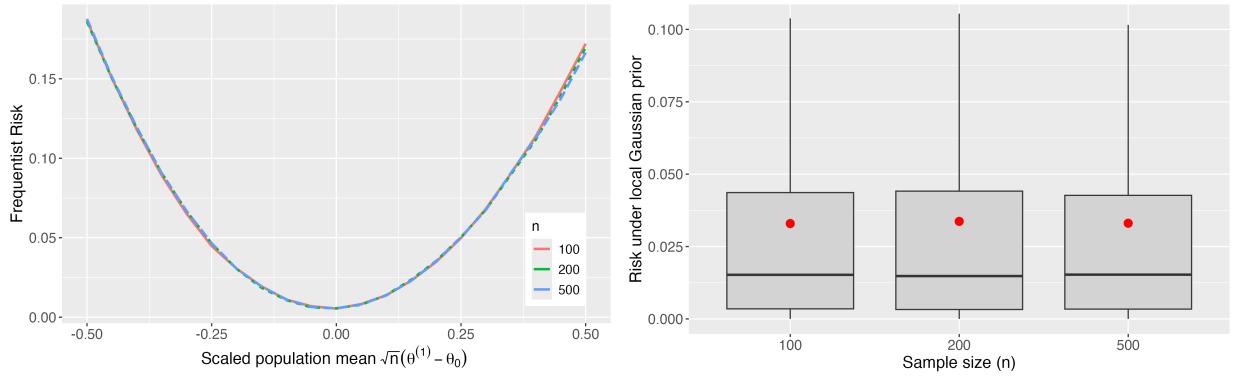


A: Frequentist risk profile

B: Bayes risk vs n

Note: Panel A shows the frequentist risk profiles for different values of n , when the data is obtained through UCB. Panel B shows how the realized risk changes with n under the prior $(h^{(1)}, h^{(0)}) \sim \text{i.i.d } \mathcal{N}(0, 0.04)$. The red points denote the Bayes risk under that prior.

FIGURE 4.1. Point estimation: two-armed UCB



A: Frequentist risk profile

B: Bayes risk vs n

Note: Panel A shows the frequentist risk profiles for different values of n , when the data is obtained through Thompson Sampling with a Beta(1, 1) prior. Panel B shows how the realized risk changes with n under the prior $(h^{(1)}, h^{(0)}) \sim \text{i.i.d } \mathcal{N}(0, 0.04)$. The red points denote the Bayes risk under that prior.

FIGURE 4.2. Point estimation: two-armed Thompson Sampling

5. APPLICATION 2: EQUIVALENCE OF IN-SAMPLE REGRET

In-sample regret measures the difference between the (welfare) performance of the best possible action—which is unknown—and a chosen policy, evaluated on the same dataset used to select the policy. As we show below, it follows from Theorem 1 that the in-sample regret from any sequence of policy rules $\{\pi_{n,j}\}_j$ can be asymptotically matched by that in the Gaussian diffusion limit experiment.⁶

Let $\mu_{n,a}(\mathbf{h}) := \mathbb{E}_{n,\mathbf{h}}[Y_i^{(a)}]$ denote the average reward corresponding to treatment a when the local parameter is \mathbf{h} . Following Hirano and Porter (2023) and Adusumilli (2025a), the reference parameter is chosen such that $\mu_{n,a}(\mathbf{0}) = 0$. The

⁶See Appendix D for extensions to out-of-sample regret, also known as simple regret.

frequentist regret of any $\pi_n \equiv \{\pi_{n,j}\}_j$ is given by

$$W_n(\mathbf{h}) = \sqrt{n} \left\{ \max_a \mu_{n,a}(\mathbf{h}) - \sum_a \mu_{n,a}(\mathbf{h}) \mathbb{E}_{n,\mathbf{h}}[q_{n,a}(1)] \right\}.$$

Analogously, in the limit experiment, the frequentist regret is given by

$$W(\mathbf{h}) = \max_a \dot{\mu}_a^\top \mathbf{h} - \sum_a \dot{\mu}_a^\top \mathbf{h} \mathbb{E}_{\mathbf{h}}[q_a(1)],$$

where $\dot{\mu}_a(\cdot)$ is defined in the assumption below:

Assumption 3. *There exists $\dot{\mu}_a \in \mathbb{R}^d$ and $\delta_n \rightarrow 0$ such that $\sqrt{n}\mu_{n,a}(\mathbf{h}) = \dot{\mu}_a^\top \mathbf{h} + \delta_n |\mathbf{h}|^2 \forall \mathbf{h}$.*

Theorem 3. *Suppose Assumptions 1-3 hold. Then, for any sequence of policies $\{\pi_{n,j}\}_j$ inducing the regret function $W_n(\mathbf{h})$, there exists a limit experiment $\{x_a(\cdot), q_a(\cdot)\}_a$ with regret function $W(\mathbf{h})$ such that $W_n(\mathbf{h}) \rightarrow W(\mathbf{h})$ for each \mathbf{h} .*

Recall from the measurability requirement on $q_a(\cdot)$ that the allocation process at any time t needs to be adapted to the sample paths of $\{x_a(s); s \leq t\}_a$. Theorem 3 thus implies that the regret profile of any policy can be asymptotically matched by one that depends only on the sample paths of $\{x_a(s); s \leq t\}_a$.

Previous work by this author (Adusumilli, 2025a) characterized the optimal Bayes and minimax risks in this setting, and showed that they can be attained by a sequence of policy rules that depend on just $\{x_a(t), q_a(t)\}_a$, i.e., the past values of these variables are not relevant. Theorem 3 is more general, in that it applies to arbitrary sequences of policy rules, but it makes use of a larger information set that includes the entire sample paths of $\{x_a(s); s \leq t\}_a$ until time t .

6. APPLICATION 3: E-PROCESSES AND ANYTIME-VALID INFERENCE

Anytime-valid inference refers to statistical procedures that maintain valid error control (e.g., Type I error or confidence coverage) uniformly over time, without compromising inference guarantees. A central tool in this framework is the e-process, a nonnegative stochastic process that starts at 1 and is a super-martingale (i.e., its expectations is always less than or equal to 1) under the null hypothesis. Much of the existing literature on anytime-valid inference has focused on sequential

experiments with fixed sampling strategies, where the only adaptive element is the stopping rule. Here, we extend the notion of an e-process to the more general setting of multi-treatment adaptive experiments.

Formally, we analyze e-processes and anytime-valid inference for tests of the form $H_0 : \boldsymbol{\theta} \in \Theta_0$ vs $H_1 : \boldsymbol{\theta} \in \Theta_1$. We assume that the reference parameter $\boldsymbol{\theta}_0 := (\theta_0^{(1)}, \theta_0^{(0)})$ from Section 2 always lies in the null region Θ_0 . As in Section 4, we index each $\boldsymbol{\theta}$ by $\mathbf{h} \in \mathbb{R}^d \times \mathbb{R}^d$ such that $\boldsymbol{\theta} = \boldsymbol{\theta}_0 + \mathbf{h}/\sqrt{n}$. Denote the set of all \mathbf{h} representing $\boldsymbol{\theta} \in \Theta_0$ by \mathcal{H}_0 and those representing $\boldsymbol{\theta} \in \Theta_1$ by \mathcal{H}_1 . We restrict attention to ‘asymptotically stable’ hypothesis testing problems, wherein the regions $\mathcal{H}_0, \mathcal{H}_1$ do not change with n .

6.1. E-processes in multi-treatment adaptive experiments. Let $\mathcal{G}_{n,\gamma_1,\gamma_0}$ denote the filtration introduced in Section 2.2.1. The e-process for multi-treatment adaptive experiments is formally defined as follows:

Definition 2. An e-process, $\varepsilon_n(q_1, q_0)$, for testing $H_0 : \mathbf{h} \in \mathcal{H}_0$ is a non-negative stochastic process indexed by $q_1, q_0 \in [0, 1]^2$, such that:

- (i) It is \mathcal{G}_{n,q_1,q_0} -adapted at any given (q_1, q_0) ; and
- (ii) For any empirical allocation process $\{q_{n,a}(\cdot)\}_a$,

$$\mathbb{E}_{n,\mathbf{h}} [\varepsilon_n(q_{n,1}(t), q_{n,0}(t))] \leq 1 \quad \forall \mathbf{h} \in \mathcal{H}_0, \forall t \in [0, 1]. \quad (6.1)$$

Definition 2 generalizes the the usual notion of an e-process to a multi-indexed super-martingale, where the indices are the treatment allocation proportions. In this framework, the value of $\varepsilon_n(\cdot)$ depends on the trajectory of the empirical allocation process, and the super-martingale property holds across all possible empirical allocation processes, i.e., all possible adaptive experiments. A central feature of the e-process is that it is algorithm-free: it remains a valid supermartingale at any time point of any adaptive experiment.

Notably, Definition 2 does not require t to be a stopping time. Optimal stopping can be incorporated by introducing a designated “default” treatment, such that assigning units to this treatment is equivalent to halting the experiment. For notational simplicity, we focus on the two-treatment case, though the extension to more arms is conceptually straightforward. If there were optimal stopping,

the interpretation of n would, however, change; it then no longer denotes a fixed sample size, but instead serves as a normalization constant. In particular, it associates the time index t to the closeness of alternatives being considered: if the aim is to analyze performance against local alternatives of the form $\boldsymbol{\theta}_0 + \mathbf{h}/\sqrt{n}$, then t should represent the number of observations collected in units of n .

E-processes serve as dynamic measures of evidence against the null hypothesis H_0 . At any point (q_0, q_1) , the value $\varepsilon_n(q_0, q_1)$ can be interpreted as the current payoff from wagering one unit against H_0 . This interpretation holds uniformly over time and across different adaptive experiments. Moreover, taking the inverse defines a p-process $p_n(q_1, q_0) := 1/\varepsilon_n(q_1, q_0)$, which yields an anytime-valid p-value. Specifically, for any $\mathbf{h} \in \mathcal{H}_0$, empirical allocation $\{q_{n,a}(\cdot)\}_a$, and time t ,

$$\begin{aligned}\mathbb{P}_{n,\mathbf{h}}(p_n(q_{n,1}(t), q_{n,0}(t)) \leq \alpha) &= \mathbb{P}_{n,\mathbf{h}}(\varepsilon_n(q_{n,1}(t), q_{n,0}(t)) \geq 1/\alpha) \\ &\leq \alpha \mathbb{E}_{n,\mathbf{h}}[\varepsilon_n(q_{n,1}(t), q_{n,0}(t))] \leq \alpha.\end{aligned}$$

Thus, $p_n(q_{n,1}(t), q_{n,0}(t))$ is a valid classical p -value at any time-point of any adaptive experiment.

6.2. GRO, mGRO and REGROW. Following Ramdas et al. (2023) and Grünwald et al. (2024), a common approach to evaluating e-processes in stopping-time experiments is through the Growth Rate Optimality (GRO) criterion. GRO assesses the quality of an e-process $\varepsilon_n(\cdot)$ based on its expected log-growth under an alternative hypothesis $\mathbf{h} \in \mathcal{H}_1$. We extend this criterion to the multi-treatment adaptive setting by defining the GRO score

$$R_n(\varepsilon_n; \mathbf{h}, \{q_{n,a}(t)\}_a) = \mathbb{E}_{n,\mathbf{h}}[\ln \varepsilon_n(q_{n,1}(t), q_{n,0}(t))]; \quad \mathbf{h} \in \mathcal{H}_1, \quad (6.2)$$

where $\{q_{n,a}(t)\}_a$ denotes the empirical allocation process at time t . The GRO score thus depends jointly on the e-process, the alternative \mathbf{h} , and the adaptive experiment employed (as indexed by the empirical allocation process).

We say that an e-process $\varepsilon_n(\cdot)$ uniformly dominates another process $\varepsilon'_n(\cdot)$ in terms of GRO if

$$R_n(\varepsilon_n; \mathbf{h}, \{q_{n,a}(t)\}_a) \geq R_n(\varepsilon'_n; \mathbf{h}, \{q_{n,a}(t)\}_a)$$

for all $\mathbf{h} \in \mathcal{H}_1$, at all time points t , and for all possible experiments, i.e., all possible empirical allocation processes $\{q_{n,a}(t)\}_a$. This notion of uniform GRO dominance is quite strong and, in general, no single e-process may achieve it. One way to relax this requirement is to instead use the mixture-GRO (mGRO) criterion, which averages the GRO score using a prior, or weight function, $w(\cdot)$ over \mathcal{H}_1 :

$$R_n(\varepsilon_n; w(\cdot), \{q_{n,a}(t)\}_a) = \int \mathbb{E}_{n,\mathbf{h}} [\ln \varepsilon_n(q_{n,1}(t), q_{n,0}(t))] dw(\mathbf{h}).$$

The mGRO criterion ranks e-processes in terms of their average performance over plausible alternatives, rather than requiring dominance for every possible alternative.

An alternative criterion, following Grünwald et al. (2024), is the REGROW (RElative GRoWth Optimality in Worst case) score:

$$\begin{aligned} \mathcal{R}_n(\varepsilon_n; \{q_{n,a}(t)\}_a) \\ = \inf_{\mathbf{h} \in \mathcal{H}_1} \left\{ \mathbb{E}_{n,\mathbf{h}} [\ln \varepsilon_n(q_{n,1}(t), q_{n,0}(t))] - \mathbb{E}_{n,\mathbf{h}} \left[\ln \frac{d\mathbb{P}_{n,\mathbf{h}}}{d\mathbb{P}_{n,0}}(q_{n,1}(t), q_{n,0}(t)) \right] \right\}. \end{aligned}$$

REGROW measures the (negative of the) worst-case GRO-regret, where GRO-regret is defined as the difference between the GRO value of $\varepsilon_n(\cdot)$ and the GRO value of the log-likelihood ratio process corresponding to a specific alternative $\mathbf{h}_1 \in \mathcal{H}_1$. The latter is the ideal, i.e., uniformly GRO optimal, e-process for testing $H_0 : \mathbf{h} = 0$ versus $H_1 : \mathbf{h} = \mathbf{h}_1$. REGROW thus benchmarks the performance of $\varepsilon_n(\cdot)$ against the optimal e-process that would be achievable if \mathbf{h}_1 were known in advance. A higher REGROW score indicates more robust performance across alternatives.

6.3. Local asymptotics and e-processes in the limit experiment. GRO and its variants, mGRO and REGROW, offer an alternative paradigm to the classical power criterion for hypothesis testing. However, determining and computing optimal e-processes becomes considerably more challenging when \mathcal{H}_0 is composite or when the REGROW criterion is used. Closed-form expressions are only known for certain parametric families $\mathbb{P}_{n,h}$.

In order to circumvent this complexity in the fixed n setting, we propose employing a local-asymptotic criterion. Accordingly, we relax the definition of e-processes

given previously, and call a sequence of non-negative, \mathcal{G}_{n,q_1,q_0} -adapted stochastic processes, $\varepsilon_n(\cdot)$, *asymptotic e-processes* if

$$\limsup_n \mathbb{E}_{n,\mathbf{h}} [\varepsilon_n(q_{n,1}(t), q_{n,0}(t))] \leq 1 \quad \forall \mathbf{h} \in \mathcal{H}_0, \quad \forall t \in [0, 1],$$

and for all possible empirical allocation processes $\{q_{n,a}(\cdot)\}_a$.

We then define an equivalent notion of an e-process in the limit experiment. Recall that the limit experiment is characterized by Gaussian process signals $z_a(q) := I_a^{1/2} h^{(a)} q + W_a(q)$. The e-process in the limit experiment is a \mathcal{G}_{q_1,q_0} -adapted stochastic process designed for testing $H_0 : \mathbf{h} \in \mathcal{H}_0$ versus $H_1 : \mathbf{h} \in \mathcal{H}_1$. The formal definition is given below. First, a bit of terminology: let \mathcal{Q} denote the collection of allocation processes in the limit experiment that are weak limits of some sequence of empirical allocation processes $\{q_{n,a}(\cdot)\}_a$.

Definition 3. An e-process, $\varepsilon(\cdot, \cdot)$, for testing $H_0 : \mathbf{h} \in \mathcal{H}_0$ in the limit experiment is a non-negative stochastic process indexed by $q_1, q_0 \in [0, 1]^2$, such that:

- (i) It is \mathcal{G}_{q_1,q_0} -adapted at any given (q_1, q_0) ; and
- (ii) For any allocation process $\{q_a(\cdot)\}_a \in \mathcal{Q}$,

$$\mathbb{E}_{\mathbf{h}}[\varepsilon(q_1(t), q_0(t))] \leq 1 \quad \forall \mathbf{h} \in \mathcal{H}_0, \quad \forall t \in [0, 1]. \quad (6.3)$$

Relative to Definition 2, the above definition restricts the set of allocation processes to \mathcal{Q} . The rationale behind this is mainly technical: it allows us to avoid dealing with $\{q_a(\cdot)\}_a$ that are not weak limit points of sequences of empirical allocation processes in the actual experiment. Whether \mathcal{Q} , in fact, includes all possible allocation processes is currently unknown (to this author).

The GRO, mGRO and REGROW criteria in the limit experiment retain the same form as (6.2), except that $\mathbb{E}_{n,\mathbf{h}}[\cdot]$ is replaced by $\mathbb{E}_{\mathbf{h}}[\cdot]$, e.g., the GRO criterion in the limit experiment becomes

$$R(\varepsilon; \mathbf{h}, \{q_a(t)\}_a) = \mathbb{E}_{\mathbf{h}} [\ln \varepsilon(q_1(t), q_0(t))]; \quad \mathbf{h} \in \mathcal{H}_1.$$

6.4. Representation theorems for e-processes. We derive an asymptotic representation theorem for e-processes, which establishes that for any sequence of asymptotic e-processes, there exists a dominating e-process in the limit experiment

with respect to the GRO criterion and its variants. This is based on the following regularity conditions:

Assumption 4. *As functions of $\{q_a\}_a$, the sequence $\{z_{n,a}(\cdot), \varepsilon_n(\cdot, \cdot)\}_a$ converges weakly under $\mathbb{P}_{n,0}$. Furthermore, for each t and each possible weakly convergent sequence of empirical processes $\{q_{n,a}(\cdot)\}_a$, the sequence $\{\ln \varepsilon_n(q_{n,1}(t), q_{n,0}(t))\}_n$ is uniformly integrable with respect to each element in $\{\mathbb{P}_{n,h}\}_{h \in \mathcal{H}_1}$.*

The first part of Assumption 4 is an analogue of Assumption 2. It restricts the class of asymptotic e-processes by requiring them to be asymptotically equicontinuous (so that they have a weak limit). The second part of Assumption 4 is an additional regularity condition ensuring that the GRO scores are asymptotically convergent under the alternative hypotheses. Both properties will need to be verified on a case-by-case basis. For an example, see Appendix E.1.

Theorem 4. *Suppose Assumptions 1 and 4 hold. Then, for any sequence of asymptotic e-processes $\varepsilon_n(\cdot)$, there exists an e-process $\varepsilon(\cdot)$ in the limit experiment—depending only on $\{z_a(q_a), q_a\}_a$ —such that*

$$\limsup_n R_n(\varepsilon_n; \mathbf{h}, \{q_{n,a}(t)\}_a) \leq R(\varepsilon; \mathbf{h}, \{q_a(t)\}_a)$$

for all $\mathbf{h} \in \mathcal{H}_1$, all $t \in [0, 1]$, and all sequences of empirical allocation processes, $\{q_{n,a}(\cdot)\}_a$, converging to some allocation process $\{q_a(\cdot)\}_a$ in the limit experiment.

An important implication of Theorem 4 is that any asymptotically optimal e-process need depend only on $\{z_a(q_a)\}_a$, in addition to its index $\{q_a\}_a$. Thus, at any time-point t , the set of sufficient statistics for anytime-valid inference is again $\{x_a(t), q_a(t)\}_a$, and the past values of these processes are not relevant for inference.

The extension to the mGRO criterion is a straightforward consequence of Theorem 4 and the monotone convergence theorem.

Corollary 3. *Suppose Assumptions 1 and 4 hold. Assume further that there exists $g(\mathbf{h}) \in [0, \infty)$ satisfying $\int g(\mathbf{h}) dw(\mathbf{h}) < \infty$ and $R_n(\varepsilon_n; \mathbf{h}, \{q_{n,a}(t)\}_a) \geq -g(\mathbf{h})$ for all \mathbf{h} , all allocation processes $\{q_{n,a}(t)\}_n$, and all sufficiently large n . Then, for any sequence of asymptotic e-processes $\varepsilon_n(\cdot)$, there exists an e-process $\varepsilon(\cdot)$ in the limit*

experiment—depending only on $\{z_a(q_a), q_a\}_a$ —such that

$$\limsup_n R_n(\varepsilon_n; w, \{q_{n,a}(t)\}_a) \leq R(\varepsilon; w, \{q_a(t)\}_a)$$

for all $t \in [0, 1]$, and all sequences of empirical allocation processes, $\{q_{n,a}(\cdot)\}_a$, converging to some allocation process $\{q_a(\cdot)\}_a$ in the limit experiment.

In most cases, the function $g(\cdot)$ in the statement of Corollary 3 can be set to 0; see Appendix E.1, for instance.

Analysis of the REGROW criterion requires an additional assumption:

Assumption 5. For any time t , $\mathbf{h}_1 \in \mathcal{H}_1$, and any sequence of empirical allocation processes $\{q_{n,a}(\cdot)\}_a$ weakly converging to some $\{q_{n,a}(\cdot)\}_a \in \mathcal{Q}$, we have that $\mathbb{E}_{n,\mathbf{h}_1} \left[\ln \frac{d\mathbb{P}_{n,\mathbf{h}_1}}{d\mathbb{P}_{n,0}}(q_{n,1}(t), q_{n,0}(t)) \right]$ converges to $\mathbb{E}_{\mathbf{h}_1} \left[\ln \frac{d\mathbb{P}_{\mathbf{h}_1}}{d\mathbb{P}_0}(q_1(t), q_0(t)) \right]$.

The assumption states that the KL-divergences, $\text{KL}(\mathbb{P}_{n,\mathbf{h}_1} || \mathbb{P}_{n,0})$, in the original experiment converge asymptotically to KL-divergences, $\text{KL}(\mathbb{P}_{\mathbf{h}_1} || \mathbb{P}_0)$, in the limit experiment. This involves restrictions on the parametric models allowed.

Corollary 4. Suppose Assumptions 1, 4 and 5 hold. Then, for any sequence of asymptotic e-processes $\varepsilon_n(\cdot)$, there exists an e-process $\varepsilon(\cdot)$ in the limit experiment—depending only on $\{z_a(q_a), q_a\}_a$ —such that

$$\limsup_n \mathcal{R}_n(\varepsilon_n; \{q_{n,a}(t)\}_a) \leq \mathcal{R}(\varepsilon; \{q_a(t)\}_a)$$

for all $t \in [0, 1]$, and all sequences of empirical allocation processes, $\{q_{n,a}(\cdot)\}_a$, converging to some allocation process $\{q_a(\cdot)\}_a$ in the limit experiment.

6.5. Applying the representation theorems. In what follows, we simplify matters by assuming that the null hypothesis \mathcal{H}_0 is a singleton, consisting solely of the reference parameter $\boldsymbol{\theta}_0$.⁷ Theorem 4 and Corollaries 3, 4 establish asymptotic upper bounds on the GRO, mGRO and REGROW criteria. These bounds are obtained by optimizing the respective criteria within the limit experiment.

⁷When there is only adaptive stopping, compound nulls are typically addressed in anytime-valid inference using the method of reverse information projection (see Ramdas et al., 2023, for a survey). Extending this approach to the adaptive sampling setting is more involved and left for future work. See, however, Appendix E for the simpler case of testing parameters corresponding to a single treatment arm.

A natural question is whether the restriction $\{q_a(\cdot)\}_a \in \mathcal{Q}$, imposed in the definition of the limiting e-process (Definition 3), affects the outcome of the optimization. In practice, it does not. The general approach to optimizing these criteria involves constructing, for each fixed allocation process $\{q_a(\cdot)\}_a$ and time t , an e-value: an \mathcal{I}_t -measurable random variable ε satisfying

$$\mathbb{E}_{\mathbf{h}}[\varepsilon] \leq 1 \quad \text{for all } \mathbf{h} \in \mathcal{H}_0.$$

We then identify the point-wise optimal e-value, $\varepsilon_{q_1(t), q_0(t)}^*$, that maximizes the desired criterion (GRO, mGRO, or REGROW) at the given allocation point $\{q_a(t)\}_a$. The GRO, mGRO or REGROW value of $\varepsilon_{q_1(t), q_0(t)}^*$ furnishes a sharp upper bound—uniformly over all possible e-processes—for the corresponding criterion evaluated at that allocation.

The final step is to determine whether these pointwise-optimal e-values can be coherently combined, or “strung together”, into a full e-process satisfying Definition 3. If such a construction is not possible, it usually implies that a globally optimal e-process—one that simultaneously achieves pointwise optimality at all allocation points—does not exist.

6.5.1. *mGRO optimality.* As a first illustration of this approach, consider the mGRO criterion. For a given $\{q_a(t)\}_a$, the mGRO optimal e-value in the limit experiment is

$$\varepsilon_{q_1(t), q_0(t)}^* = \int \exp \sum_a \left\{ h^{(a)\top} I_a^{1/2} z_a(q_a(t)) - \frac{q_a(t)}{2} h^{(a)\top} I_a h^{(a)} \right\} dw(\mathbf{h}).$$

Importantly, the form of $\varepsilon_{q_1(t), q_0(t)}^*$ does not change with $\{q_a(t)\}_a$, implying that the optimal e-process $\varepsilon^*(\cdot, \cdot)$ can be constructed as $\varepsilon^*(q_1, q_0) = \varepsilon_{q_1, q_0}^*$. It is straightforward to verify that the resulting e-process satisfies (6.3) using Lemma 1 and standard martingale arguments.

Replacing $\{z_a(\cdot)\}_a$ with $\{z_{n,a}(\cdot)\}_a$ yields the asymptotically mGRO-optimal e-process:

$$\varepsilon_n^*(q_1, q_0) = \int \exp \sum_a \left\{ h^{(a)\top} I_a^{1/2} z_{n,a}(q_a) - \frac{q_a}{2} h^{(a)\top} I_a h^{(a)} \right\} dw(\mathbf{h}). \quad (6.4)$$

Appendix E.1 describes primitive conditions under which $\varepsilon_n^*(q_1, q_0)$ satisfies the requirements for Corollary 3. Essentially, we need $w(\cdot)$ to be a sub-Gaussian distribution and $\psi_a(Y_i^{(a)})$ to have finite $2 + p$ moments, for some $p > 0$.

6.5.2. REGROW optimality. Grünwald et al. (2024) show that the REGROW-optimal e-value, $\bar{\varepsilon}_{q_1(t), q_0(t)}^*$, coincides with the mGRO-optimal e-value for a specific, least-favorable weighting function $w_{q_1(t), q_0(t)}^*(\cdot)$. The authors also show that this weighting function is obtained as the solution to

$$w_{q_1(t), q_0(t)}^* = \arg \max_{w \in \Delta(\mathcal{H}_1)} \mathbb{E}_{\mathbf{h} \sim w} [\text{KL}_{q_1(t), q_0(t)} (\mathbb{P}_{\mathbf{h}} \parallel \mathbb{P}_w)], \quad (6.5)$$

where $\Delta(\mathcal{H}_1)$ denotes the set of all probability measures supported on \mathcal{H}_1 ,

$$\mathbb{P}_w(\cdot) := \int \mathbb{P}_{\mathbf{h}}(\cdot) dw(\mathbf{h}),$$

and $\text{KL}_{q_1(t), q_0(t)} (\mathbb{P}_{\mathbf{h}} \parallel \mathbb{P}_w)$ represents the KL divergence between $\mathbb{P}_{\mathbf{h}}, \mathbb{P}_w$ when these probability measures are restricted to the filtration $\mathcal{I}_t \equiv \mathcal{G}_{q_1(t), q_0(t)}$.⁸

Grünwald et al. (2024) further demonstrate that the optimized value of the objective in (6.5) provides an upper bound on the REGROW criterion $\mathcal{R}(\varepsilon; \{q_a(t)\}_a)$:

$$\sup_{\varepsilon} \mathcal{R}(\varepsilon; \{q_a(t)\}_a) \leq \mathbb{E}_{\mathbf{h} \sim w_{q_1(t), q_0(t)}^*} [\text{KL}_{q_1(t), q_0(t)} (\mathbb{P}_{\mathbf{h}} \parallel \mathbb{P}_w)].$$

Replacing the $q_1(t), q_0(t)$ subscripts with (\cdot) for ease of notation, note that $w_{(\cdot)}^*$ can be alternatively characterized as:

$$w_{(\cdot)}^* = \arg \max_{w \in \Delta(\mathcal{H}_1)} \text{KL}_{(\cdot)} (p_{\mathbf{h}} \cdot w \parallel p_w \cdot w) = \arg \max_{w \in \Delta(\mathcal{H}_1)} I_{(\cdot)} (w; p_w), \quad (6.6)$$

where $p_{\mathbf{h}}, p_w$ denote the densities of $\mathbb{P}_{\mathbf{h}}, \mathbb{P}_w$ with respect to some dominating measure, and $I_{(\cdot)}(\cdot; \cdot)$ represents mutual information under the restricted filtration. The optimization problem (6.6) has a natural information-theoretic interpretation. Consider an information transmission channel with input \mathbf{h} and output $\{z_a(s); s \leq q_a(t)\}_a$. Then, the quantity $\sup_w I_{(\cdot)}(w; p_w)$ corresponds to the channel capacity, and $w_{(\cdot)}^*$ represents the optimal signal distribution.

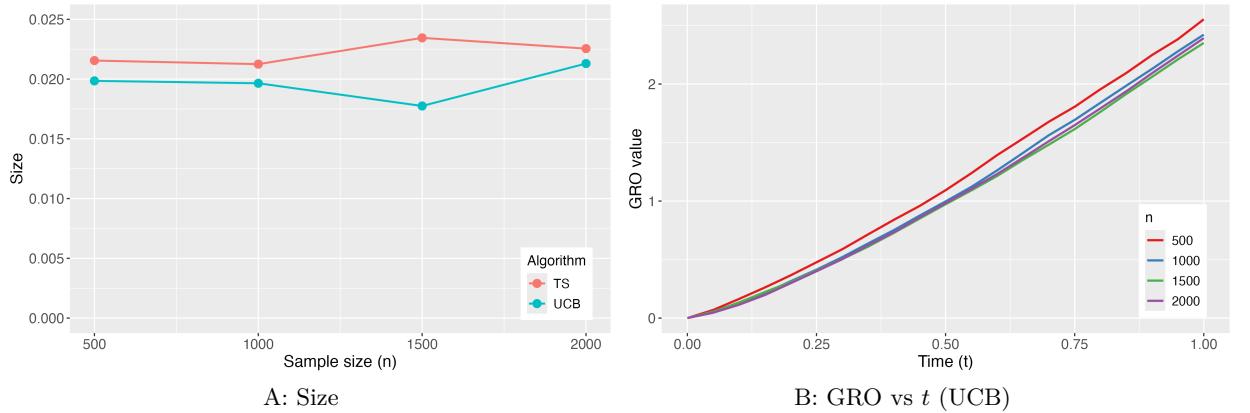
⁸For instance, the restriction of $\mathbb{P}_{\mathbf{h}}$ to \mathcal{I}_t is the probability measure induced by the sample paths of $z_a(s) = I_a^{1/2} h^{(a)} s + W_a(s)$ between 0 and $q_a(t)$.

Crucially, the least-favorable distribution $w_{(\cdot)}^*$ depends on both the chosen allocation process and the structure of the alternative hypothesis set \mathcal{H}_1 . It is essential that \mathcal{H}_1 be compact—without it, the channel capacity (and hence the REGROW value) becomes infinite. Because of the dependence on the allocation process, a globally optimal REGROW e-process does not exist. In practice, one must fix a specific allocation process $\{q_a(t)\}_a$ and an alternative region \mathcal{H}_1 , from which an optimal weighting function $w_{(\cdot)}^*$ can be derived, e.g., by employing the Blahut-Arimoto algorithm (see, Arimoto, 1972; Blahut, 1972). The resulting e-process is thus *locally REGROW optimal* relative to the chosen design and alternative set (it would also be globally mGRO optimal relative to the least favorable distribution $w^*(\cdot)$). Appendix E provides an example of such a locally REGROW optimal e-process, constructed to be optimal against fixed values of $\{q_a\}_a$.

6.6. Illustrative example (contd.) We revisit again the illustrative example from Section 2.1, now focusing on constructing an anytime-valid test of the null hypothesis $H_0 : \theta^{(1)} = 0.1$ against the two-sided alternative $H_1 : \theta^{(1)} \neq 0.1$. We take the reference parameter vector to be $\boldsymbol{\theta}_0 = (0.1, 0.1)$ —as it is the only reference parameter that induces non-trivial asymptotic limits—and consider local alternatives of the form $\boldsymbol{\theta} = \boldsymbol{\theta}_0 + \mathbf{h}/\sqrt{n}$.

Although the null is composite due to the unrestricted nature of $h^{(0)}$, Appendix E shows that it is without loss of generality to ignore observations from arm 0. This dimensionality reduction allows us to construct optimal e-processes using only the data from arm 1, leveraging the techniques developed in Section 6.5.

We employ the mGRO criterion with the weighting function $w_1(\cdot) \sim \mathcal{N}(0, \nu^2)$, where $\nu^2 = 1$. This corresponds to a $\mathcal{N}(0.1, 1/n)$ prior over $\theta^{(1)}$. We examine sample sizes $n \in \{500, 1000, 1500, 2000\}$. For instance, when $n = 1000$, this prior places 95% of the mass within the credible interval $[0.06, 0.14]$ for $\theta^{(1)}$. As noted earlier, in real-world scenarios, the difference in click-through rates between various variants of a website is typically less than 1%. Hence, our chosen weighting aligns well with realistic alternative values of $\theta^{(1)}$.



Note: Panel A displays the uniform over time finite sample size of the e-process in equation (6.7) at different values of n . Panel B plots the evolution of GRO value of this test over time under UCB, at the local alternative $(h^{(1)}, h^{(0)}) = (1/\sqrt{n}, 0)$.

FIGURE 6.1. Anytime-valid inference

Under this setup, the asymptotically optimal mGRO e-process (from equation 6.4) takes the explicit form:

$$\varepsilon_n^*(q_1) = \frac{1}{\sqrt{1 + q_1 I_1 \nu^2}} \exp \left[\frac{I_1^2 \nu^2 \left\{ \sum_{j=1}^{\lfloor nq_1 \rfloor} (Y_j^{(1)} - 0.1) \right\}^2}{2n(1 + q_1 I_1 \nu^2)} \right], \quad (6.7)$$

where $I_1 := [\theta_0^{(1)}(1 - \theta_0^{(1)})]^{-1}$ denotes the Fisher information.

Panel A of Figure 6.1 reports the uniform-over-time finite-sample size of this e-process under Thompson Sampling and UCB allocation rules, calculated as⁹

$$\sup_{h^{(0)}} P_{n,(0,h^{(0)})} \left(\max_t \varepsilon_n^*(q_{n,1}(t)) \geq 20 \right),$$

The critical value of 20 implies that the p-process conversion (from the e-process) targets an anytime valid size of 5%. However, the size under specific policies can be smaller and it indeed turns out that the test is conservative for the policies considered—a behavior that is expected, since the validity of the e-process applies to all adaptive algorithms and not just Thompson Sampling or UCB.

Panel B plots the evolution of the GRO value over time for this e-process under a local alternative $\mathbf{h} \equiv (h^{(1)}, h^{(0)}) = (1/\sqrt{n}, 0)$, for varying sample sizes, all under UCB (see Appendix E for equivalent results under Thompson Sampling). The resulting curves exhibit remarkable stability across n , indicating that the asymptotic approximation is already accurate at such small sample sizes as $n = 500$.

⁹The least-favorable configuration for $h^{(0)}$ appears to be $-\infty$, corresponding to $\theta^{(0)} = 0$.

7. CONCLUSION

This article introduces a continuous-time formalism for analyzing fully adaptive experiments. The formalism is based on the notion of allocation processes, also introduced in this article. We show that any empirical allocation process, as induced by some policy rule, converges weakly to a corresponding allocation process in a limit experiment governed by Gaussian diffusions. The various applications, ranging from point-estimation to anytime-valid inference illustrate the utility and generality of this framework.

Beyond these applications, the continuous-time formulation offers a powerful tool for addressing design problems in adaptive experimentation. Though not based on the results reported in this article, prior work by this author has applied the continuous-time framework to derive optimal algorithms for bandit experiments and costly sampling problems (Adusumilli 2025a,b). The current results provide an easy-to-use and generalizable template for transferring optimal designs from the limit experiment back to the finite-sample setting. Looking ahead, we expect this approach to be broadly useful in addressing a range of open questions in adaptive experimentation—e.g., in deriving optimal strategies for best-arm identification with multiple treatments, or in designing adaptive experiments in strategic environments involving interactions between an experimenter and a regulator.

REFERENCES

- K. Adusumilli, “Optimal tests following sequential experiments,” *arXiv preprint arXiv:2305.00403*, 2023.
- , “Risk and optimal policies in bandit experiments,” *Econometrica*, vol. 93, no. 3, pp. 1003–1029, 2025.
- , “How to sample and when to stop sampling: The generalised Wald problem and minimax policies,” *The Review of Economic Studies*, 2025.
- S. Arimoto, “An algorithm for computing the capacity of arbitrary discrete memoryless channels,” *IEEE Transactions on Information Theory*, vol. 18, no. 1, pp. 14–20, 1972.
- K. J. Arrow, D. Blackwell, and M. A. Girshick, “Bayes and minimax solutions of sequential decision problems,” *Econometrica*, pp. 213–244, 1949.

- D. Blackwell, “Equivalent comparisons of experiments,” *The Annals of Mathematical Statistics*, pp. 265–272, 1953.
- R. Blahut, “Computation of channel capacity and rate-distortion functions,” *IEEE Transactions on Information Theory*, vol. 18, no. 4, pp. 460–473, 1972.
- D. Bouneffouf and I. Rish, “A survey on practical applications of multi-armed and contextual bandits,” *arXiv preprint arXiv:1904.10040*, 2019.
- J. Chapman, E. Snowberg, S. W. Wang, and C. Camerer, “Dynamically optimized sequential experimentation (DOSE) for estimating economic preference parameters,” National Bureau of Economic Research, Tech. Rep., 2024.
- A. Deng, Y. Xu, R. Kohavi, and T. Walker, “Improving the sensitivity of online controlled experiments by utilizing pre-experiment data,” in *Proceedings of the sixth ACM international conference on Web search and data mining*, 2013, pp. 123–132.
- L. Fan and P. W. Glynn, “Diffusion approximations for Thompson Sampling,” *arXiv preprint arXiv:2105.09232*, 2021.
- F. Finan and D. Pouzo, “Reinforcing RCTs with multiple priors while learning about external validity,” National Bureau of Economic Research, Tech. Rep., 2022.
- P. Grünwald, R. de Heide, and W. Koolen, “Safe testing,” *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 2024.
- E. Häusler and H. Luschgy, *Stable convergence and stable limit theorems*. Springer, 2015, vol. 74.
- K. Hirano and J. R. Porter, “Asymptotic representations for sequential decisions, adaptive experiments, and batched bandits,” *arXiv preprint arXiv:2302.03117*, 2023.
- A. Kalvit and A. Zeevi, “A closer look at the worst-case behavior of multi-armed bandit algorithms,” *Advances in Neural Information Processing Systems*, vol. 34, pp. 8807–8819, 2021.
- I. Karatzas and S. Shreve, *Brownian motion and stochastic calculus*. Springer Science & Business Media, 2012, vol. 113.
- M. Kasy and A. Sautmann, “Adaptive treatment assignment in experiments for policy choice,” *Econometrica*, vol. 89, no. 1, pp. 113–132, 2021.

- X. Kuang and S. Wager, “Weak signal asymptotics for sequentially randomized experiments,” *Management Science*, vol. 70, no. 10, pp. 7024–7041, 2024.
- T. Lattimore and C. Szepesvári, *Bandit algorithms*. Cambridge University Press, 2020.
- L. Le Cam, “A reduction theorem for certain sequential experiments. II,” *The Annals of Statistics*, vol. 7, no. 4, pp. 847 – 859, 1979.
- L. M. Le Cam, *Asymptotic methods in statistical theory*. Springer-Verlag, 1986.
- A. Ramdas, P. Grünwald, V. Vovk, and G. Shafer, “Game-theoretic statistics and safe anytime-valid inference,” *Statistical Science*, vol. 38, no. 4, pp. 576–601, 2023.
- J. G. Smith, “The information capacity of amplitude- and variance-constrained scalar Gaussian channels,” *Information and Control*, vol. 18, no. 3, pp. 203–219, Apr. 1971.
- A. W. van der Vaart and J. Wellner, *Weak convergence and empirical processes: With applications to statistics*. Springer Science & Business Media, 1996.
- A. Wald, *Sequential analysis*. John Wiley, 1947.
- G. Wassmer and W. Brannath, *Group sequential and confirmatory adaptive designs in clinical trials*. Springer, 2016, vol. 301.

APPENDIX A. PROOFS OF LEMMA 1 AND THEOREM 1

A.1. Proof of Lemma 1. We start by showing that $x_1(t)$ is an \mathcal{I}_t -martingale; the claim for $x_0(t)$ follows analogously.

Fix any $t_1, t_2 \in [0, 1]$ such that $t_2 \geq t_1$, and define $\tau_1 := q_1(t_1), \tau_2 := q_1(t_2)$. Also, for each $\gamma_1 \geq 0$, define $\mathcal{H}_{\gamma_1} := \mathcal{G}_{\gamma_1, 1}$. By the definition of the allocation process, the event $\{q_1(t) \leq \gamma_1\}$ is \mathcal{H}_{γ_1} measurable, under any given t . Hence, τ_1, τ_2 are both $\{\mathcal{H}_{\gamma_1}\}_{\gamma_1 \geq 0}$ -adapted stopping times. It is easily verified from the definition of $\mathcal{G}_{\gamma_1, \gamma_0}$ that $z_1(\gamma_1)$ is a Wiener process with respect to \mathcal{H}_{γ_1} . Since $\tau_2 \geq \tau_1$ almost surely (a.s.), due to the almost sure monotonicity of $q(\cdot)$, it follows by the optional sampling theorem that $\mathbb{E}[z(\tau_2)|\mathcal{H}_{\tau_1}] = z(\tau_1)$ a.s. In other words, $\mathbb{E}[z_1(q_1(t_2))|\mathcal{H}_{q_1(t_1)}] = z_1(q_1(t_1))$ a.s., i.e., $\mathbb{E}[x_1(t_2)|\mathcal{H}_{q_1(t_1)}] = x_1(t_1)$. Since

$\mathcal{I}_t \equiv \mathcal{G}_{q_1(t), q_0(t)} \subseteq \mathcal{G}_{q_1(t), 1} \equiv \mathcal{H}_{q_1(t)}$ for any t , the tower property of conditional expectations then implies $\mathbb{E}[x_1(t_2)|\mathcal{I}_{t_1}] = x_1(t_1)$ a.s. But t_1, t_2 were arbitrary, so the claim follows.

To show that the quadratic variation of $x_1(t)$ is $q_1(t)$, we start by observing that $G(\gamma_1) := z_1^2(\gamma_1) - \gamma_1$ is a \mathcal{H}_{γ_1} -martingale by the well known properties of Wiener processes. Then, by similar arguments as above, we have that $\mathbb{E}[G(\tau_2)|\mathcal{H}_{\tau_1}] = G(\tau_1)$ almost surely, and it thereby follows, also by similar arguments as before, that $\mathbb{E}[x_1^2(t_2) - q_1(t_2)|\mathcal{I}_{t_1}] = x_1^2(t_1) - q_1(t_1)$ almost surely. Hence, $x_1^2(t) - q_1(t)$ is an \mathcal{I}_t -martingale. This proves that $q_1(t)$ is the quadratic variation of $x_1(t)$. An analogous argument shows that $q_0(t)$ is the quadratic variation of $x_0(t)$.

A.2. Proof of Theorem 1. We can informally outline the proof as follows: We employ dyadic approximations to discretize the empirical allocation processes $\{q_{n,a}(\cdot)\}_a$ across both their support (i.e., time) and range. The resulting discrete processes are shown to converge in distribution under $\mathbb{P}_{n,0}$ to a limit (discrete) process that is a function of $(z_1(\cdot), z_0(\cdot), U)$, where $z_1(\cdot), z_0(\cdot)$ are independent d -dimensional Wiener processes, and $U \sim \text{Uniform}[0, 1]$ represents exogenous randomization.

Next, we demonstrate that as the discretization becomes arbitrarily fine, the intermediate limit processes converge to a continuous-time allocation process satisfying the criteria in Definition 1. The theory of stable convergence (Häusler and Luschgy, 2015) plays a key role here in ensuring that the informational/measurability constraints satisfied by the empirical allocation processes are preserved during the transition to the continuous-time (and range) limit.

The proof proceeds in the following steps:

Step 1 (Convergence under dyadic approximations):

We discretize time into dyadic sets $\mathcal{D}_m = \{t_k : k = 0, \dots, 2^m\}$, where $t_k = k2^{-m}$. Denote

$$q_{n,a,k}^{(m)} := q_{n,a}(t_k). \quad (\text{A.1})$$

We then employ a further dyadic discretization $\mathcal{D}_l \equiv \{\eta_0\} \cup \{\eta_l : l = 1, \dots, 2^L\}$ of the range, $[0, 1]$, of the empirical allocation process, $q_{n,1}(\cdot)$, where $\eta_0 = 0$, $\eta_l := l2^{-L}$ and $2^L := \bar{c}2^m$ for some natural number $\bar{c} > 1$. Subsequently, we

approximate each of the random variables $q_{n,1,k}^{(m)}$ by

$$q_{n,1,k}^{(m,L)} = \sum_{l=1}^{2^L} \eta_l \mathbb{I}\{\eta_{l-1} < q_{n,1,k}^{(m)} \leq \eta_l\}. \quad (\text{A.2})$$

Let $\phi_{n,k,l}$ denote the indicator functions $\mathbb{I}\{q_{n,1,k}^{(m,L)} = \eta_l\}$. The random variables $\{\phi_{n,k,l}\}_{k,l}$ are tight, as are the processes $z_{n,a}(\cdot)$, the latter due to standard results in empirical process theory. Hence, by Prohorov's theorem, there exists a subsequence, $\{n_k\}_k$, represented as n for simplicity, such that

$$\begin{pmatrix} \{\phi_{n,k,l}\}_{k,l} \\ z_{n,1}(\cdot) \\ z_{n,0}(\cdot) \end{pmatrix} \xrightarrow[\mathbb{P}_{n,0}]{} \begin{pmatrix} \{\tilde{\phi}_{k,l}\}_{k,l} \\ z_1(\cdot) \\ z_0(\cdot) \end{pmatrix}, \quad (\text{A.3})$$

where $z_1(\cdot), z_0(\cdot)$ are independent d -dimensional Wiener processes. Denote

$$\tilde{q}_{1,k}^{(m,L)} := \sum_{l=1}^{2^L} \eta_l \tilde{\phi}_{k,l}, \quad \tilde{q}_{0,k}^{(m,L)} := t_k - \tilde{q}_{1,k}^{(m,L)},$$

$$s_{n,a,l} := z_{n,a}(\eta_l) - z_{n,a}(\eta_{l-1}), \text{ and}$$

$$s_{a,l} := z_a(\eta_l) - z_a(\eta_{l-1}).$$

Lemma 2 in Appendix C shows that we can construct versions of $\tilde{q}_{a,k}^{(m,L)}$, denoted $q_{a,k}^{(m,L)}$, that satisfy the following conditions:

C1: $q_{1,k}^{(m,L)} + q_{0,k}^{(m,L)} = t_k$ for all k .

C2: For each k , $\{q_{1,k}^{(m,L)} \leq \eta_l, q_{0,k}^{(m,L)} \leq \eta_{l'}\}$ is

$$\sigma\{U_1, \dots, U_k, \{s_{1,j}\}_{j \leq l+1}, \{s_{0,j}\}_{j \leq l'+1}\}$$

measurable for all $l + l' \geq k\bar{c}$, where U_1, \dots, U_k are uniform random variables independent of $z_1(\cdot), z_0(\cdot)$.

The random variables U_1, \dots, U_{2^m} can be subsumed into a single $U \sim \text{Uniform}[0, 1]$. Also, define $l(\gamma) = \inf\{l : \eta_l > \gamma\}$. Then, we can rewrite the first part of this condition as: $\{q_{1,k}^{(m,L)} \leq \gamma_1, q_{0,k}^{(m,L)} \leq \gamma_0\}$ is

$$\sigma\{\{s_{1,j}\}_{j \leq l(\gamma_1)+1}, \{s_{0,j}\}_{j \leq l(\gamma_0)+1}, U\} \subseteq \bar{\mathcal{F}}_{\gamma_1+2^{-L}}^{(1)} \vee \bar{\mathcal{F}}_{\gamma_0+2^{-L}}^{(0)} \vee \sigma(U)$$

measurable for any $\gamma_1, \gamma_0 \in [0, 1]$ such that $\gamma_1 + \gamma_0 \geq t_k$.

C3: Letting \sim denote equivalence in distributions,

$$\left(\{s_{a,l}\}_{a,l}, \{\tilde{q}_{1,\tilde{k}}^{(m,L)}\}_{\tilde{k}=1}^k \right) \sim \left(\{s_{a,l}\}_{a,l}, \{q_{1,\tilde{k}}^{(m,L)}\}_{\tilde{k}=1}^k \right) \text{ for each } k. \quad (\text{A.4})$$

Equations (A.3) and (A.4) imply

$$\left(\{s_{n,a,l}\}_{a,l}, \{q_{n,a,k}^{(m,L)}\}_{a,k} \right) \xrightarrow[\mathbb{P}_{n,0}]{} \left(\{s_{a,l}\}_{a,l}, \{q_{a,k}^{(m,L)}\}_{a,k} \right).$$

As a particular consequence of the above and the continuous mapping theorem, if we define $x_{n,a,k}^{(m,L)} := z_{n,a}(q_{n,a,k}^{(m,L)})$ and $x_{n,a}^{(m,L)} := z_a(q_{a,k}^{(m,L)})$, then

$$\left\{ x_{n,a,k}^{(m,L)}, q_{n,a,k}^{(m,L)} \right\}_{a,k} \xrightarrow[\mathbb{P}_{n,0}]{} \left\{ x_{a,k}^{(m,L)}, q_{a,k}^{(m,L)} \right\}_{a,k}. \quad (\text{A.5})$$

Note that, by construction, $q_{n,a,k}^{(m,L)} \leq t_k$ and $q_{n,a,k}^{(m)} \leq q_{n,a,k'}^{(m)}$ for all k and $k' \geq k$. Hence (A.5) implies—by the properties of weak convergence—that we also have $q_{a,k}^{(m,L)} \leq t_k$ and $q_{a,k}^{(m,L)} \leq q_{a,k'}^{(m,L)}$ almost surely, for all k and $k' \geq k$.

Step 2 (Taking $L \rightarrow \infty$):

Equations (A.3)-(A.5) apply under any fixed L . In fact, as $z_1(\cdot), z_0(\cdot), U$ do not depend on L , and $\{q_{a,k}^{(m,L)}\}_{a,k}$ is a measurable function of these quantities by Condition C2, we can construct versions of $\{z_1(\cdot), z_0(\cdot), \{q_{a,k}^{(m,L)}\}_{a,k}\}$ that lie in the same probability space and where $z_1(\cdot), z_0(\cdot), U$ are the same quantities across L . Since $q_{a,k}^{(m,L)}$ is tight (it takes values in $[0, 1]$), by Prohorov's theorem, there exists a sequence $L_j \rightarrow 0$ and some random variables $\{q_{a,k}^{(m)}\}_k$ such that

$$\left(U, z_1(\cdot), z_0(\cdot), \left\{ q_{a,k}^{(m,L_j)} \right\}_{a,k} \right) \xrightarrow{d} \left(U, z_1(\cdot), z_0(\cdot), \{q_{a,k}^{(m)}\}_{a,k} \right) \text{ as } j \rightarrow \infty. \quad (\text{A.6})$$

Recall from the end of Step 1 that $q_{a,k}^{(m,L_j)} \leq t_k$, $q_{a,k}^{(m,L_j)} \leq q_{a,k'}^{(m,L_j)}$ and $q_{1,k}^{(m,L_j)} + q_{0,k}^{(m,L_j)} = t_k$ almost surely, for all k and $k' \geq k$. Equation (A.6) then implies $q_{a,k}^{(m)} \leq t_k$, $q_{a,k}^{(m)} \leq q_{a,k'}^{(m)}$ and $q_{1,k}^{(m)} + q_{0,k}^{(m)} = t_k$ almost surely, for all k and $k' \geq k$.

Define

$$\begin{aligned} q_a^{(m,L)}(t) &= \sum_{k=0}^{2^m-1} q_{a,k}^{(m,L)} \mathbb{I}\{t_k \leq t < t_{k+1}\}, \text{ and} \\ q_a^{(m)}(t) &= \sum_{k=0}^{2^m-1} q_{a,k}^{(m)} \mathbb{I}\{t_k \leq t < t_{k+1}\}. \end{aligned} \quad (\text{A.7})$$

It then follows from (A.6) that

$$\left\{ U, z_a(\cdot), q_a^{(m, L_j)}(\cdot) \right\}_a \xrightarrow{d} \left\{ U, z_a(\cdot), q_a^{(m)}(\cdot) \right\}_a \text{ as } j \rightarrow \infty. \quad (\text{A.8})$$

Another implication of (A.6) is that

$$\left\{ x_{a,k}^{(m, L_j)}, q_{a,k}^{(m, L_j)} \right\}_{a,k} \xrightarrow{d} \left\{ x_{a,k}^{(m)}, q_{a,k}^{(m)} \right\}_{a,k} \text{ as } j \rightarrow \infty, \quad (\text{A.9})$$

where

$$x_{a,k}^{(m)} := z_a(q_{a,k}^{(m)}).$$

We conclude this step by deriving some limit approximations for $q_{n,a,k}^{(m)}$ (defined in A.2) and $x_{n,a,k}^{(m)} := z_{n,a}(q_{n,a,k}^{(m)})$. The dyadic discretization $q_{n,a,k}^{(m, L_j)}$ of $q_{n,a,k}^{(m)}$ is such that $q_{n,a,k}^{(m, L_j)} \downarrow q_{n,a,k}^{(m)}$ and

$$\sup_{n,k} \left| q_{n,a,k}^{(m)} - q_{n,a,k}^{(m, L_j)} \right| \leq 2^{-L_j} \rightarrow 0 \text{ as } j \rightarrow \infty. \quad (\text{A.10})$$

Recall the definition $x_{n,a,k}^{(m,L)} := z_{n,a}(q_{n,a,k}^{(m,L)})$ from Step 1. For every $\epsilon > 0$ and $a \in \{0, 1\}$, (A.10) implies

$$\begin{aligned} \limsup_{n \rightarrow \infty} \mathbb{P}_{n,0} \left(\sup_k \left| x_{n,a,k}^{(m)} - x_{n,a,k}^{(m, L_j)} \right| > \epsilon \right) &\leq \limsup_{n \rightarrow \infty} \mathbb{P}_{n,0} \left(\sup_{q \in [0,1], \delta \in [0, 2^{-L_j}]} |z_{n,a}(q + \delta) - z_{n,a}(q)| > \epsilon \right) \\ &\rightarrow 0 \text{ as } j \rightarrow \infty, \end{aligned} \quad (\text{A.11})$$

where the second step follows from Karatzas and Shreve (2012, Lemma 2.4.19).

Combining (A.5), (A.9) and (A.11), we conclude

$$\left\{ x_{n,a,k}^{(m)}, q_{n,a,k}^{(m)} \right\}_{a,k} \xrightarrow[\mathbb{P}_{n,0}]{} \left\{ x_{a,k}^{(m)}, q_{a,k}^{(m)} \right\}_{a,k}. \quad (\text{A.12})$$

Step 3 (Taking $m \rightarrow \infty$):

Equation (A.8) applies for any fixed m . Therefore the construction in Step 2 can be applied for each m , giving rise to a sequence of processes $\{z_a(\cdot), q_a^{(m)}(\cdot), U\}_a$.

By construction, $\sum_a q_a^{(m)}(t) = t \forall t$ almost surely since, as shown in Step 2, $q_{1,k}^{(m)} + q_{0,k}^{(m)} = t_k \forall k$ almost surely. Furthermore, as $q_{a,k}^{(m)} \leq q_{a,k+1}^{(m)}$ for all k almost surely (as also shown in Step 2), it follows that $q_1^{(m)}(\cdot), q_0^{(m)}(\cdot)$ are also almost surely monotone. We now claim that the sequence $\{q_a^{(m)}(\cdot)\}_{m=1}^\infty$ is stochastically equicontinuous. Recall the definition of $\{q_{n,1,k}^{(m)}\}_k$ from (A.1) and observe that by

the structure of $q_{n,1}(\cdot)$,

$$\sup_k \left| q_{n,1,k+1}^{(m)} - q_{n,1,k}^{(m)} \right| \leq 2^{-m} + n^{-1}.$$

In view of (A.12), the above implies

$$\mathbb{P} \left(\sup_k \left| q_{1,k+1}^{(m)} - q_{1,k}^{(m)} \right| > 2^{-m} \right) = 0$$

for each m . Consequently, from the definition of $q_1^{(m)}(\cdot)$ in (A.7), it follows that for any $\delta > 0$,

$$\mathbb{P} \left(\sup_t \left| q_1^{(m)}(t + \delta) - q_1^{(m)}(t) \right| > \delta + 2^{-m} \right) = 0. \quad (\text{A.13})$$

This implies $\{q_1^{(m)}(\cdot)\}_{m=1}^\infty$ is stochastically equicontinuous. Stochastic equicontinuity of $\{q_0^{(m)}(\cdot)\}_m$ also follows since $q_0^{(m)}(t) = t - q_1^{(m)}(t)$ almost surely.

As it is stochastically equicontinuous, the sequence $\{q_1^{(m)}(\cdot), q_0^{(m)}(\cdot)\}_{m=1}^\infty$ is tight. Combined with the tightness of $\{z_a(\cdot)\}_a, U$, we conclude that the joint $\{z_a(\cdot), U, q_a^{(m)}(\cdot)\}_a$ is also tight. Then, by Prohorov's theorem, there exists a subsequence $\{m_k\}_{k=1}^\infty$, represented as $\{m\}$ without loss of generality, such that

$$\{z_a(\cdot), U, q_a^{(m)}(\cdot)\}_a \xrightarrow{d} \{z_a(\cdot), U, q_a(\cdot)\}_a. \quad (\text{A.14})$$

Step 4 (Existence of an allocation process satisfying A.14):

We now show there exists a version of $q_a(\cdot)$, defined on a suitably constructed probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$, that is a valid allocation process.

First, (A.14) and $\sum_a q_a^{(m)}(t) = t \forall t$ (as shown in Step 3) implies $\sum_a q_a(t) = t \forall t$ almost surely. Second, $q_a(\cdot)$ is almost surely monotone as it is the weak limit of almost surely monotone processes $\{q_a^{(m)}(\cdot)\}_a \in D[0, 1]^2$ and the set of monotone functions is closed under the Skorokhod topology. It thus remains to construct a version of $q_a(\cdot)$ that satisfies the measurability requirement of Definition 1.

By (A.8) and (A.14), there exists a sequence $\{(m_j, L_j)\}_{j=1}^\infty$ with $(m_j, L_j) \rightarrow (\infty, \infty)$, under which¹⁰

$$\{z_a(\cdot), U, q_a^{(m_j, L_j)}(\cdot)\}_a \xrightarrow{d} \{z_a(\cdot), U, q_a(\cdot)\}_a \text{ as } j \rightarrow \infty. \quad (\text{A.15})$$

¹⁰It is possible to choose such a sequence since weak convergence can be metrized using the bounded Lipschitz metric.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ represent the canonical probability space corresponding to $\{z_a(\cdot)\}_a, U$. Specifically, we set $\Omega \equiv C^d[0, 1] \times C^d[0, 1] \times [0, 1]$, $\mathcal{F} \equiv \mathcal{B}(\Omega)$ and $\mathbb{P} \equiv \mathcal{W}^d \otimes \mathcal{W}^d \otimes \lambda$, where $C^d[0, 1]$ represents the space of continuous functions from $[0, 1]$ to \mathbb{R}^d , \mathcal{W}^d denotes the d -dimensional Wiener measure on $C^d[0, 1]$, and λ is the Lebesgue measure on $[0, 1]$. Then, (A.15) implies, by the properties of stable convergence, that¹¹

$$\left\{q_a^{(m_j, L_j)}(\cdot)\right\}_a \xrightarrow[\mathcal{F}]{\text{stably}} \{q_a(\cdot)\}_a. \quad (\text{A.16})$$

Set $\gamma := (\gamma_1, \gamma_0)$, $\mathbf{Q}_j(t) := (q_1^{(m_j, L_j)}(t), q_0^{(m_j, L_j)}(t))$ and $\mathbf{Q}(t) := (q_1(t), q_0(t))$. Since $\mathbf{Q}_j(\cdot)$ takes values on the Skorokhod space $D[0, 1]^2$ —which is Polish—stable convergence $\mathbf{Q}_j(\cdot) \xrightarrow[\mathcal{F}]{\text{stably}} \mathbf{Q}(\cdot)$ implies there exists a Markov kernel, $K(\omega, dy)$, from Ω to $D[0, 1]^2$ that acts as the limit version of the conditional distribution of \mathbf{Q}_j given \mathcal{F} .¹² We can then construct a measurable representation of $\mathbf{Q}(\cdot)$ on the extended probability space, $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}}) \equiv (\Omega \times [0, 1], \mathcal{F} \otimes \mathcal{B}[0, 1], \mathbb{P} \otimes \lambda)$, such that $K(\omega, dy)$ represents the conditional probability of $\mathbf{Q}(\cdot)$ given \mathcal{F} . In essence, the extended probability space augments the underlying set of variables $\{z_a(\cdot)\}_a, U$ with another exogenous randomization $V \sim \text{Uniform}[0, 1]$. By the usual properties of stable convergence, (A.15) continues to hold for this representation of $\mathbf{Q}(\cdot)$.

Let $\mathcal{F}_\gamma^{(a)} \subseteq \mathcal{F}$ denote the right-continuous filtration generated by the sample paths of $z_a(\cdot)$ between 0 and γ , and take $\mathcal{G}_{\gamma_1, \gamma_0} \subseteq \mathcal{F}$ to be the augmentation of $\mathcal{F}_{\gamma_1}^{(1)} \vee \mathcal{F}_{\gamma_0}^{(0)} \vee \sigma(U)$ with respect to $(\Omega, \mathcal{F}, \mathbb{P})$. In the extended probability space, this gives rise to the extended filtration $\bar{\mathcal{G}}_{\gamma_1, \gamma_0} \equiv \mathcal{G}_{\gamma_1, \gamma_0} \vee \sigma(V) \subseteq \bar{\mathcal{F}}$. Note that the filtration $\bar{\mathcal{G}}_{\gamma_1, \gamma_0}$ inherits the right-continuous and augmented nature of $\mathcal{G}_{\gamma_1, \gamma_0}$. We now argue that $\{\mathbf{Q}(t) \leq \gamma\} \equiv \{q_1(t) \leq \gamma_1, q_0(t) \leq \gamma_0\}$ is $\bar{\mathcal{G}}_{\gamma_1, \gamma_0}$ measurable for each $\gamma_1, \gamma_0, t \in [0, 1]$ such that $\gamma_1 + \gamma_0 \geq t$.

For any $\epsilon > 0$ and $\mathbf{u} = (u_1, u_0) \in [0, 1]^2$, let $\phi_{\epsilon, \gamma}(\mathbf{u})$ denote a smoothed version of $\mathbb{I}\{u_1 \leq \gamma_1, u_0 \leq \gamma_0\}$, defined as

$$\phi_{\epsilon, \gamma}(\mathbf{u}) = \begin{cases} 1 & \text{if } u_1 \leq \gamma_1, u_0 \leq \gamma_0 \\ 0 & \text{if } u_1 \geq \gamma_1 + \epsilon \text{ or } u_0 \geq \gamma_0 + \epsilon \\ \text{linear decay} & \text{otherwise.} \end{cases}$$

¹¹See Häusler and Luschgy (2015) for a textbook treatment of stable convergence.

¹²Formally, $K(\omega, dy)$ is such that $\mathbb{E}[f(\mathbf{Q}_j)|\mathcal{F}] \xrightarrow{P} \int f(y)K(\cdot, dy)$ for all bounded continuous $f(\cdot)$.

Given a fixed value of j , let $k \in \{0, \dots, 2^{m_j}\}$, $t_k := k2^{-m_j}$ and $k(t) := \max\{k : t \geq t_k\}$. By the definition of $q_a^{(m_j, L_j)}(\cdot)$, along with Condition C2 from Step 1, $\{\mathbf{Q}_j(t) \leq \boldsymbol{\gamma} + \epsilon'\} \equiv \{q_{1, k(t)}^{(m_j, L_j)} \leq \gamma_1 + \epsilon', q_{0, k(t)}^{(m_j, L_j)} \leq \gamma_0 + \epsilon'\}$ is

$$\mathcal{G}_{\gamma_1 + \epsilon' + 2^{-L_j}, \gamma_0 + \epsilon' + 2^{-L_j}} \subseteq \mathcal{G}_{\gamma_1 + 2\epsilon', \gamma_0 + 2\epsilon'}$$

measurable for all $\epsilon' \in [0, \epsilon]$ and j sufficiently large. This implies, by the definition of $\phi_{\epsilon, \boldsymbol{\gamma}}(\cdot)$, that $\phi_{\epsilon, \boldsymbol{\gamma}}(\mathbf{Q}_j(t))$ is $\mathcal{G}_{\gamma_1 + 2\epsilon, \gamma_0 + 2\epsilon}$ -measurable. Hence, for any bounded \mathcal{F} -measurable random variable W that is independent of $\mathcal{G}_{\gamma_1 + 2\epsilon, \gamma_0 + 2\epsilon}$ (i.e., W relies only on Wiener process increments “after” $\mathcal{G}_{\gamma_1 + 2\epsilon, \gamma_0 + 2\epsilon}$), we must have

$$\mathbb{E}[\phi_{\epsilon, \boldsymbol{\gamma}}(\mathbf{Q}_j(t))W] = \mathbb{E}[\phi_{\epsilon, \boldsymbol{\gamma}}(\mathbf{Q}_j(t))]\mathbb{E}[W]. \quad (\text{A.17})$$

The definition of stable convergence states that

$$\mathbb{E}[f(\mathbf{Q}_j(t))Z] \rightarrow \bar{\mathbb{E}}[f(\mathbf{Q}(t))Z] \text{ as } j \rightarrow \infty$$

for any bounded continuous function $f(\cdot)$, and any bounded \mathcal{F} -measurable random variable Z . Applying the above to the factorization (A.17) with $f = \phi_{\epsilon, \boldsymbol{\gamma}}$ and $Z = \{1, W\}$, we get

$$\bar{\mathbb{E}}[\phi_{\epsilon, \boldsymbol{\gamma}}(\mathbf{Q}(t))W] = \bar{\mathbb{E}}[\phi_{\epsilon, \boldsymbol{\gamma}}(\mathbf{Q}(t))]\mathbb{E}[W]. \quad (\text{A.18})$$

(A.18) holds for any bounded random variable W independent of $\mathcal{G}_{\gamma_1 + 2\epsilon, \gamma_0 + 2\epsilon}$ in the original probability space. But space of all such W is equivalent to the space of all bounded random variables independent of $\bar{\mathcal{G}}_{\gamma_1 + 2\epsilon, \gamma_0 + 2\epsilon}$ in the extended probability space. Hence, (A.18) implies $\phi_{\epsilon, \boldsymbol{\gamma}}(\mathbf{Q}(t))$ is independent of the “future” Wiener-process noise relative to $\bar{\mathcal{G}}_{\gamma_1 + 2\epsilon, \gamma_0 + 2\epsilon}$, and is therefore $\bar{\mathcal{G}}_{\gamma_1 + 2\epsilon, \gamma_0 + 2\epsilon}$ -measurable.

We are interested in the event $E \equiv \{\mathbf{Q}(t) \leq \boldsymbol{\gamma}\}$. Notice, from the definition of $\phi_{\epsilon, \boldsymbol{\gamma}}(\cdot)$, that $\mathbb{I}_E = \lim_{\epsilon \downarrow 0} \phi_{\epsilon, \boldsymbol{\gamma}}(\mathbf{Q}(t))$ point-wise for each $\omega \in \bar{\Omega}$. But as shown earlier, $\phi_{\epsilon, \boldsymbol{\gamma}}(\mathbf{Q}(t))$ is $\bar{\mathcal{G}}_{\gamma_1 + 2\epsilon, \gamma_0 + 2\epsilon}$ -measurable; consequently, \mathbb{I}_E must be $\cap_{\epsilon \downarrow 0} \bar{\mathcal{G}}_{\gamma_1 + 2\epsilon, \gamma_0 + 2\epsilon} \equiv \bar{\mathcal{G}}_{\gamma_1, \gamma_0}$ -measurable, where the equivalence is due to the right continuity of the filtrations.

This concludes the existence of a valid allocation process $\{q_a(\cdot)\}_a$ —defined on an extended probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ —that satisfies (A.14). But the additional

$V \sim \text{Uniform}[0,1]$ randomization employed in the definition of $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ can be combined with U to form a new $\text{Uniform}[0, 1]$ random variable \bar{U} . So the measurability requirement of Definition 1 is satisfied by taking the relevant probability space to be $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$.

Step 5 (Completing the proof):

Define $x_a^{(m)}(t) = z_a(q_a^{(m)}(t))$ and $x_a(t) = z_a(q_a(t))$. By construction,

$$x_a^{(m)}(t) = \sum_{k=0}^{2^m-1} x_{a,k}^{(m)} \{t_k \leq t < t_{k+1}\},$$

where $x_{a,k}^{(m)} = z_a(q_{a,k}^{(m)})$, as defined earlier in Step 2. Since $q_a(\cdot)$ is the limit of stochastically equicontinuous processes, it has almost surely continuous sample paths. Combined with (A.14), this implies

$$\{x_a^{(m)}(\cdot), U, q_a^{(m)}(\cdot)\}_a \xrightarrow{d} \{x_a(\cdot), U, q_a(\cdot)\}_a \text{ as } m \rightarrow \infty. \quad (\text{A.19})$$

By the properties of weak convergence (see, e.g., van der Vaart and Wellner 1996, Chapter 1.12), part (iii) of Theorem 1 follows if we show

$$\mathbb{E}_{n,0} [f(x_{n,1}(\cdot), x_{n,0}(\cdot), q_{n,1}(\cdot), q_{n,0}(\cdot))] \rightarrow \mathbb{E} [f(x_1(\cdot), x_0(\cdot), q_1(\cdot), q_0(\cdot))] \quad (\text{A.20})$$

for all bounded Lipschitz continuous $f(\cdot)$.

Fix a value of m and construct dyadic approximations for $x_{n,a}(\cdot), q_{n,a}(\cdot)$ of the form

$$\begin{aligned} x_{n,a}^{(m)}(t) &= \sum_{k=0}^{2^m-1} x_{n,a}(t_k) \mathbb{I}\{t_k \leq t < t_{k+1}\} \equiv \sum_{k=0}^{2^m-1} x_{n,a,k}^{(m)} \mathbb{I}\{t_k \leq t < t_{k+1}\}, \\ q_{n,a}^{(m)}(t) &= \sum_{k=0}^{2^m-1} q_{n,a}(t_k) \mathbb{I}\{t_k \leq t < t_{k+1}\} \equiv \sum_{k=0}^{2^m-1} q_{n,a,k}^{(m)} \mathbb{I}\{t_k \leq t < t_{k+1}\}. \end{aligned}$$

In what follows, let $S_n := \{x_{n,a}(\cdot), q_{n,a}(\cdot)\}_a$, $S_n^{(m)} := \{x_{n,a}^{(m)}(\cdot), q_{n,a}^{(m)}(\cdot)\}_a$, $S := \{x_a(\cdot), q_a(\cdot)\}_a$ and $S^{(m)} := \{x_a^{(m)}(\cdot), q_a^{(m)}(\cdot)\}_a$. We can then decompose

$$\begin{aligned} \mathbb{E}_{n,0} [f(S_n)] - \mathbb{E} [f(S)] &= \left\{ \mathbb{E} [f(S^{(m)})] - \mathbb{E} [f(S)] \right\} + \left\{ \mathbb{E}_{n,0} [f(S_n^{(m)})] - \mathbb{E} [f(S^{(m)})] \right\} \\ &\quad + \left\{ \mathbb{E}_{n,0} [f(S_n)] - \mathbb{E}_{n,0} [f(S_n^{(m)})] \right\} \\ &:= T_1^{(m)} + T_{2n}^{(m)} + T_{3n}^{(m)}. \end{aligned}$$

By (A.19), $|T_1^{(m)}| \rightarrow 0$ as $m \rightarrow \infty$. Furthermore, (A.12) implies $\lim_{n \rightarrow \infty} T_{2n}^{(m)} = 0$ for any given m . It remains to bound $T_{3n}^{(m)}$. Note that by the definition of $q_{n,a}(\cdot)$,

$$\sup_{t \in [0,1]} |q_{n,a}(t + 2^{-m}) - q_{n,a}(t)| \leq 2^{-m} + n^{-1} := \delta_{m,n}$$

for any $m > 0$. Then, letting B denote the upper bound of $|f(\cdot)|$ and C the Lipschitz constant of $f(\cdot)$, we observe that for every $\epsilon > 0$, and all n sufficiently large so that $\delta_{m,n} < 2^{-(m-1)}$,

$$|T_{3n}^{(m)}| \leq C(2^{-m} + n^{-1}) + \sum_{a \in \{0,1\}} \left\{ C\epsilon + 2B \cdot \mathbb{P}_{n,0} \left(\sup_{q \in [0,1], \delta \in [0,2^{-(m-1)}]} |z_{n,a}(q + \delta) - z_{n,a}(q)| > \epsilon \right) \right\}.$$

Define

$$r_m^{(a)}(\epsilon) := \limsup_{n \rightarrow \infty} \mathbb{P}_{n,0} \left(\sup_{q \in [0,1], \delta \in [0,2^{-(m-1)}]} |z_{n,a}(q + \delta) - z_{n,a}(q)| > \epsilon \right).$$

Then, for any $\epsilon > 0$,

$$\limsup_{n \rightarrow \infty} |T_{3n}^{(m)}| \leq C2^{-m} + 2C\epsilon + 2B(r_m^{(1)}(\epsilon) + r_m^{(0)}(\epsilon)) := \bar{r}(m, \epsilon),$$

By Karatzas and Shreve (2012, Lemma 2.4.19), $\lim_{m \rightarrow \infty} r_m^{(a)}(\epsilon) = 0$ for any $\epsilon > 0$, so $\lim_{m \rightarrow \infty} \bar{r}(m, \epsilon) = 2C\epsilon$.

To conclude, we have shown that for any given m, ϵ ,

$$\limsup_{n \rightarrow \infty} |\mathbb{E}_{n,0}[f(S_n)] - \mathbb{E}[f(S)]| \leq |T_1^{(m)}| + \bar{r}(m, \epsilon). \quad (\text{A.21})$$

In view of the previous results, the right hand side of (A.21) can be made arbitrarily small by taking $m \rightarrow \infty$ and $\epsilon \rightarrow 0$. This proves (A.20).

It remains to show that $q_a(\cdot)$ is Lipschitz continuous. But this is a straightforward consequence of (A.13) and the fact that $q_a(\cdot)$ is the almost sure limit of $q_a^{(m)}(\cdot)$, see (A.15).

SUPPLEMENTARY APPENDIX

APPENDIX B. PROOFS OF THE REMAINING RESULTS

B.1. Proof of Corollary 1. Recall the quantity $\hat{\varphi}(\mathbf{h}; \gamma_1, \gamma_0)$ from Section 3. By (3.1),

$$\hat{\varphi}(\mathbf{h}; q_{n,1}(1), q_{n,0}(1)) = \sum_a \left\{ h^{(a)\top} I_a^{1/2} x_{n,a}(1) - \frac{q_{n,a}(1)}{2} h^{(a)\top} I_a h^{(a)} \right\} + o_{\mathbb{P}_{n,0}}(1).$$

Combining the above with Theorem 1 and Assumption 2 gives

$$\begin{aligned} \begin{pmatrix} \{x_{n,a}(\cdot), q_{n,a}(\cdot)\}_a \\ \hat{\varphi}(\mathbf{h}; q_{n,1}(1), q_{n,0}(1)) \end{pmatrix} &\xrightarrow[\mathbb{P}_{n,0}]{} \begin{pmatrix} \{x_a(\cdot), q_a(\cdot)\}_a \\ V \end{pmatrix}; \text{ where} \\ V &\sim \exp \sum_a \left\{ h^{(a)\top} I_a^{1/2} x_a(1) - \frac{q_a(1)}{2} h^{(a)\top} I_a h^{(a)} \right\}. \end{aligned} \quad (\text{B.1})$$

Denote

$$S(t) := \sum_a h^{(a)\top} I_a^{1/2} x_a(t)$$

and

$$M(t) := \exp \sum_a \left\{ h^{(a)\top} I_a^{1/2} x_a(t) - \frac{q_a(t)}{2} h^{(a)\top} I_a h^{(a)} \right\}.$$

By Lemma 1, $S(t)$ is an \mathcal{I}_t -martingale and its quadratic variation is given by $\sum_a \frac{q_a(t)}{2} h^{(a)\top} I_a h^{(a)}$. Hence, $M(t)$ is the stochastic/Doleans-Dade exponential of $S(t)$. As $q_a(t) \leq 1$ almost surely,

$$E \left[\exp \int_0^t \left\{ \sum_a \frac{q_a(t)}{2} h^{(a)\top} I_a h^{(a)} \right\} dt \right] \leq \exp \left\{ \sum_a \frac{1}{2} h^{(a)\top} I_a h^{(a)} \right\} < \infty.$$

Thus, Novikov's condition is satisfied and $M(t)$ is also an \mathcal{I}_t -martingale. Doob's optional sampling theorem then implies $E[V] \equiv E[M(1)] = E[M(0)] = 1$.

Since the processes, $\{x_a(\cdot), q_a(\cdot)\}_a$, are tight, their sample paths lie in a separable metric space \mathcal{D} , with an associated Borel sigma-algebra $\mathcal{B}(\mathcal{D})$. This, together with the fact that $V \geq 0$ and $\mathbb{E}[V] = 1$, implies, by a version of Le Cam's third lemma for processes (see, e.g., van der Vaart and Wellner 1996, Theorem 3.10.7), that

$$\{x_{n,a}(\cdot), q_{n,a}(\cdot)\}_a \xrightarrow[\mathbb{P}_{n,h}]{} \mathcal{L}; \text{ where } \mathcal{L}(B) := E [\mathbb{I} (\{x_a(\cdot), q_a(\cdot)\}_a \in B) V] \quad \forall B \in \mathcal{B}(\mathcal{D}).$$

But $\{x_a(\cdot), q_a(\cdot)\}_a$ is adapted to $\mathcal{I}_1 := \mathcal{G}_{q_1(1), q_0(1)}$ due to Definition 1, so, by the Girsanov theorem,

$$\begin{aligned}\mathcal{L}(B) &= E \left[\mathbb{I}(\{x_a(\cdot), q_a(\cdot)\}_a \in B) \exp \sum_a \left\{ h^{(a)\top} I_a^{1/2} z_a(q_a(1)) - \frac{q_a(1)}{2} h^{(a)\top} I_a h^{(a)} \right\} \right], \\ &= \mathbb{P}_{\boldsymbol{h}} (\{x_a(\cdot), q_a(\cdot)\}_a \in B),\end{aligned}$$

where the probability $\mathbb{P}_{\boldsymbol{h}}$ (defined in Section 3.1) is the one induced by the sample paths of $z_a(\cdot) \sim I_a^{1/2} h_a \cdot + W_a(\cdot)$, together with an exogenous randomization U . The claim therefore follows.

B.2. Proof of Theorem 2. Recall the definition $\hat{\varphi}(\boldsymbol{h}; \gamma_1, \gamma_0)$ from Section 3. By (3.1),

$$\hat{\varphi}(\boldsymbol{h}; q_{n,1}(1), q_{n,0}(1)) = \sum_a \left\{ h^{(a)\top} I_a^{1/2} x_{n,a}(1) - \frac{q_{n,a}(1)}{2} h^{(a)\top} I_a h^{(a)} \right\} + o_{\mathbb{P}_{n,0}}(1).$$

Combining the above with Theorem 1 and Assumption 2 gives

$$\hat{\varphi}(\boldsymbol{h}; q_{n,1}(1), q_{n,0}(1)) \xrightarrow[\mathbb{P}_{n,0}]{} \sum_a \left\{ h^{(a)\top} I_a^{1/2} x_a(1) - \frac{q_a(1)}{2} h^{(a)\top} I_a h^{(a)} \right\}. \quad (\text{B.2})$$

By the definition of weak regularity, the sequence $\sqrt{n}(T_n - \boldsymbol{\theta}_0)$ is tight, and consequently, so is the sequence $\sqrt{n}(T_n - \boldsymbol{\theta}(\boldsymbol{h})) = \sqrt{n}(T_n - \boldsymbol{\theta}_0) - \boldsymbol{h}$. Since the individual elements converge in distribution, it follows that the joint

$$(\sqrt{n}(T_n - \boldsymbol{\theta}(\boldsymbol{h})), \hat{\varphi}(\boldsymbol{h}; q_{n,1}(1), q_{n,0}(1)))$$

is also tight. Hence, by Prohorov's theorem, given any sequence $\{n_j\}$, there exists a further sub-sequence $\{n_{j_m}\}$ —represented as $\{n\}$ for ease of notation—and a random variable \bar{T} such that

$$\begin{aligned}\begin{pmatrix} \sqrt{n}(T_n - \boldsymbol{\theta}(\boldsymbol{h})) \\ \hat{\varphi}(\boldsymbol{h}; q_{n,1}(1), q_{n,0}(1)) \end{pmatrix} &\xrightarrow[\mathbb{P}_{n,0}]{} \begin{pmatrix} \bar{T} - \boldsymbol{h} \\ V \end{pmatrix}; \text{ where} \\ V &\sim \exp \sum_a \left\{ h^{(a)\top} I_a^{1/2} x_a(1) - \frac{q_a(1)}{2} h^{(a)\top} I_a h^{(a)} \right\}.\end{aligned} \quad (\text{B.3})$$

As in the proof of Corollary 1, $E[V] = 1$.

We now claim that

$$\sqrt{n}(T_n - \boldsymbol{\theta}(\mathbf{h})) \xrightarrow[\mathbb{P}_{n,\mathbf{h}}]{d} \mathcal{L}; \text{ where } \mathcal{L}(B) := E \left[\mathbb{I}\{\bar{T} - \mathbf{h} \in B\} V \right] \forall B \in \mathcal{B}(\mathbb{R}). \quad (\text{B.4})$$

It is clear from $V \geq 0$ and $E[V] = 1$ that \mathcal{L} is a probability measure, and that for every measurable function $f : \mathbb{R} \rightarrow \mathbb{R}$, $\int f d\mathcal{L} = E[f(\bar{T} - \mathbf{h})V]$. Furthermore, for any $f(\cdot)$ lower-semicontinuous and non-negative,

$$\begin{aligned} & \liminf \mathbb{E}_{n,\mathbf{h}} \left[f \left(\sqrt{n}(T_n - \boldsymbol{\theta}(\mathbf{h})) \right) \right] \\ & \geq \liminf \mathbb{E}_{n,0} \left[f \left(\sqrt{n}(T_n - \boldsymbol{\theta}(\mathbf{h})) \right) \frac{d\mathbb{P}_{n,\mathbf{h}}}{d\mathbb{P}_{n,0}} \right] \\ & = \liminf \mathbb{E}_{n,0} \left[f \left(\sqrt{n}(T_n - \boldsymbol{\theta}(\mathbf{h})) \right) \exp \{ \hat{\varphi}(\mathbf{h}; q_{n,1}(1), q_{n,0}(1)) \} \right] \\ & \geq E[f(\bar{T} - \mathbf{h})V]. \end{aligned} \quad (\text{B.5})$$

The equality in (B.5) follows from the law of iterated expectations since T_n is $\mathcal{I}_{n,1} := \mathcal{G}_{n,q_{n,1}(1),q_{n,0}(1)}$ measurable,

$$\frac{d\mathbb{P}_{n,\mathbf{h}}}{d\mathbb{P}_{n,0}} \equiv \frac{d\mathbb{P}_{n,\mathbf{h}}}{d\mathbb{P}_{n,0}} \left(\mathbf{y}_n^{(1)}, \mathbf{y}_n^{(0)} \right) = \exp \{ \hat{\varphi}(\mathbf{h}; 1, 1) \}$$

by definition (see Section 3.1), and

$$\mathbb{E}_{n,0} [\exp \{ \hat{\varphi}(\mathbf{h}; 1, 1) \} | \mathcal{I}_{n,1}] = \exp \{ \hat{\varphi}(\mathbf{h}; q_{n,1}(1), q_{n,0}(1)) \}$$

as the observations are iid given \mathbf{h} . The last inequality in (B.5) follows from applying the portmanteau lemma on (B.3). Applying the portmanteau lemma again, in the converse direction, on (B.5), gives (B.4).

Weak convergence, (B.4), implies that for any non-negative loss function $l(\cdot)$,

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \mathbb{E}_{n,\mathbf{h}} \left[l \left(\sqrt{n}(T_n - \boldsymbol{\theta}(\mathbf{h})) \right) \right] \\ & \geq E \left[l(\bar{T} - \mathbf{h}) \exp \sum_a \left\{ h^{(a)\top} I_a^{1/2} x_a(1) - \frac{q_a(1)}{2} h^{(a)\top} I_a h^{(a)} \right\} \right]. \end{aligned} \quad (\text{B.6})$$

Define $s := \{x_1(1), x_0(1), q_1(1), q_0(1)\}$ and $T(s) := E[\bar{T}|s]$. Since $l(\cdot)$ is convex, the conditional version of Jensen's inequality implies

$$\begin{aligned} & E \left[l(\bar{T} - \mathbf{h}) \exp \sum_a \left\{ h^{(a)\top} I_a^{1/2} x_a(1) - \frac{q_a(1)}{2} h^{(a)\top} I_a h^{(a)} \right\} \right] \\ &= E \left[E \left[l(\bar{T} - \mathbf{h}) \mid s \right] \exp \sum_a \left\{ h^{(a)\top} I_a^{1/2} x_a(1) - \frac{q_a(1)}{2} h^{(a)\top} I_a h^{(a)} \right\} \right] \\ &\geq E \left[l(T - \mathbf{h}) \exp \sum_a \left\{ h^{(a)\top} I_a^{1/2} z_a(q_a(1)) - \frac{q_a(1)}{2} h^{(a)\top} I_a h^{(a)} \right\} \right]. \end{aligned}$$

But by the Girsanov theorem, applied on the processes $\{z_a(\cdot)\}_a$, the last term is just the expectation, $\mathbb{E}_h[l(T - \mathbf{h})]$, of $l(T - \mathbf{h})$ when $x_a(t) := z_a(q_a(t))$ and $z_a(\cdot)$ is distributed as a Gaussian process with drift $I_a^{1/2} h^{(a)}$, i.e., when $z_a(q) \sim I_a^{1/2} h^{(a)} q + W_a(q)$.

B.3. Proof of Corollary 2. For any tight sequence of estimators $\{T_n\}_n$, there exists a further subsequence $\{T_{n_k}\}_k$, and an estimator T in the limit experiment such that

$$\liminf_{k \rightarrow \infty} \int R_{n_k}(T_{n_k}, \mathbf{h}) dm(\mathbf{h}) \geq \int \liminf_{k \rightarrow \infty} R_{n_k}(T_{n_k}, \mathbf{h}) dm(\mathbf{h}) \geq \int R(T, \mathbf{h}) dm(\mathbf{h}),$$

where the first inequality follows by Fatou's lemma, and the second inequality by Theorem 2.

B.4. Proof of Theorem 3. Due to Assumption 3, the claim follows if we show that $\mathbb{E}_{n,\mathbf{h}}[q_{n,a}(1)] \rightarrow \mathbb{E}_h[q_a(1)]$ for each \mathbf{h} and $a \in \{0, 1\}$. Theorem 1 and Assumption 2 gives

$$\begin{aligned} & \begin{pmatrix} q_{n,a}(1) \\ \hat{\varphi}(\mathbf{h}; q_{n,1}(1), q_{n,0}(1)) \end{pmatrix} \xrightarrow[\mathbb{P}_{n,0}]{} \begin{pmatrix} q_a(1) \\ V \end{pmatrix}; \text{ where} \\ & V \sim \exp \sum_a \left\{ h^{(a)\top} I_a^{1/2} x_a(1) - \frac{q_a(1)}{2} h^{(a)\top} I_a h^{(a)} \right\}. \end{aligned}$$

By similar arguments as in the proof of Theorem 2, the above implies

$$q_{n,a}(1) \xrightarrow[\mathbb{P}_{n,\mathbf{h}}]{d} \mathcal{L}; \text{ where } \mathcal{L}(B) := E[\mathbb{I}\{q_a(1) \in B\}V] \quad \forall B \in \mathcal{B}(\mathbb{R}).$$

Consequently,

$$\mathbb{E}_{n,\mathbf{h}}[q_{n,a}(1)] \rightarrow \mathbb{E} \left[q_a(1) e^{\sum_a \left\{ h^{(a)\top} I_a^{1/2} x_a(1) - \frac{q_a(1)}{2} h^{(a)\top} I_a h^{(a)} \right\}} \right].$$

But by the Girsanov theorem, the right hand side is just the expectation, $\mathbb{E}_{\mathbf{h}}[q_a(1)]$, of $q_a(1)$ when $x_a(t) := z_a(q_a(t))$ and $z_a(\cdot)$ is distributed as a Gaussian process with drift $I_a^{1/2} h^{(a)}$.

B.5. Proof of Theorem 4. Recall the definition of $\hat{\varphi}(\mathbf{h}; \cdot, \cdot)$ in Section 3. By (3.1),

$$\hat{\varphi}(\mathbf{h}; q_1, q_0) = \sum_a \left\{ h^{(a)\top} I_a^{1/2} z_{n,a}(q_a) - \frac{q_a}{2} h^{(a)\top} I_a h^{(a)} \right\} + o_{\mathbb{P}n,0}(1), \quad (\text{B.7})$$

uniformly over all bounded q_1, q_0 .

By Assumption 4 and (B.7),

$$\begin{pmatrix} \varepsilon_n(\cdot, \cdot) \\ \hat{\varphi}(\mathbf{h}; \cdot, \cdot) \end{pmatrix} \xrightarrow[\mathbb{P}_{n,0}]{} \begin{pmatrix} \bar{\varepsilon}(\cdot, \cdot) \\ V(\cdot, \cdot) \end{pmatrix}; \text{ where} \\ V(q_1, q_0) \sim \exp \sum_a \left\{ h^{(a)\top} I_a^{1/2} z_a(q_a) - \frac{q_a}{2} h^{(a)\top} I_a h^{(a)} \right\}.$$
(B.8)

For any given q_1, q_0 , define $\varepsilon(q_1, q_0) = E[\bar{\varepsilon}(q_1, q_0) | \{z_a(q_a)\}_a]$. Then, as a process, $\varepsilon(\cdot, \cdot)$ is \mathcal{G}_{q_1, q_0} -adapted by construction.

We now claim that $\varepsilon(\cdot, \cdot)$ is a valid e-process in the limit experiment. To this end, let $\{q_{n,a}(\cdot)\}_a$ denote an arbitrary sequence empirical allocation processes whose limit point is $\{q_a(\cdot)\}_a$. By (B.8) and Theorem 1,

$$\begin{pmatrix} \varepsilon_n(q_{n,1}(t), q_{n,0}(t)) \\ \hat{\varphi}(\mathbf{h}; q_{n,1}(t), q_{n,0}(t)) \end{pmatrix} \xrightarrow[\mathbb{P}_{n,0}]{} \begin{pmatrix} \bar{\varepsilon}(q_1(t), q_0(t)) \\ V(q_1(t), q_0(t)) \end{pmatrix}.$$

As in the proof of Theorem 2, $E[V(q_1(t), q_0(t))] = 1$. Furthermore, by the arguments as in that proof again, we also have

$$\begin{aligned} \varepsilon_n(q_{n,1}(t), q_{n,0}(t)) &\xrightarrow[\mathbb{P}_{n,\mathbf{h}}]{} \mathcal{L}; \text{ where,} \\ \mathcal{L}(B) &:= E[\mathbb{I}\{\bar{\varepsilon}(q_1(t), q_0(t)) \in B\} V(q_1(t), q_0(t))] \quad \forall B \in \mathcal{B}(\mathbb{R}). \end{aligned} \quad (\text{B.9})$$

As $\varepsilon_n(q_{n,1}(t), q_{n,0}(t))$ is sequence of non-negative random variables, the portman-teau lemma and (B.9) imply

$$\begin{aligned} & \liminf_n \mathbb{E}_{n,\mathbf{h}} [\varepsilon_n(q_{n,1}(t), q_{n,0}(t))] \\ & \geq E \left[\bar{\varepsilon}(q_1(t), q_0(t)) \exp \left\{ \sum_a \left\{ h^{(a)\top} I_a^{1/2} z_a(q_a(t)) - \frac{q_a(t)}{2} h^{(a)\top} I_a h^{(a)} \right\} \right\} \right] \end{aligned}$$

for each \mathbf{h} . Furthermore, by the law of iterated expectations,

$$\begin{aligned} & E \left[\bar{\varepsilon}(q_1(t), q_0(t)) \exp \left\{ \sum_a \left\{ h^{(a)\top} I_a^{1/2} z_a(q_a(t)) - \frac{q_a(t)}{2} h^{(a)\top} I_a h^{(a)} \right\} \right\} \right] \\ & = E \left[\varepsilon(q_1(t), q_0(t)) \exp \left\{ \sum_a \left\{ h^{(a)\top} I_a^{1/2} z_a(q_a(t)) - \frac{q_a(t)}{2} h^{(a)\top} I_a h^{(a)} \right\} \right\} \right] \\ & = \mathbb{E}_{\mathbf{h}} [\varepsilon(q_1(t), q_0(t))], \end{aligned}$$

where the last step follows by the Girsanov theorem as in the proof of Theorem 2. But $\liminf_n \mathbb{E}_{n,\mathbf{h}} [\varepsilon_n(q_{n,1}(t), q_{n,0}(t))] \leq 1$ for any $\mathbf{h} \in \mathcal{H}_0$ and $t \in [0, 1]$ by the definition of an asymptotic e-process, so we conclude by the above argument that for any allocation process $\{q_a(\cdot)\}_a \in \mathcal{Q}$,

$$\mathbb{E}_{\mathbf{h}} [\varepsilon(q_1(t), q_0(t))] \leq 1 \quad \forall \mathbf{h} \in \mathcal{H}_0, \quad t \in [0, 1].$$

Since we have previously shown that $\varepsilon(\cdot, \cdot)$ is \mathcal{G}_{q_1, q_0} -adapted, the above implies that $\varepsilon(\cdot, \cdot)$ is a valid e-process in the limit experiment (in the sense of Definition 3).

Equation (B.9) and Assumption 4 also imply that for each $\mathbf{h} \in \mathcal{H}_1$ and allocation processes $\{q_{n,a}(\cdot)\}_a$ converging to an allocation process $\{q_a(\cdot)\}_a$ in the limit experiment,

$$\lim_{n \rightarrow \infty} \mathbb{E}_{n,\mathbf{h}} [\ln \varepsilon_n(q_{n,1}(t), q_{n,0}(t))] = E [V(q_1(t), q_0(t)) \ln \bar{\varepsilon}(q_1(t), q_0(t))].$$

But by the conditional Jensen's inequality and the Girsanov theorem,

$$\begin{aligned} E [V(q_1(t), q_0(t)) \ln \bar{\varepsilon}(q_1(t), q_0(t))] & = E \left[V(q_1(t), q_0(t)) E \left[\ln \bar{\varepsilon}(q_1(t), q_0(t)) \mid \mathcal{G}_{q_1(t), q_0(t)} \right] \right] \\ & \leq E [V(q_1(t), q_0(t)) \ln \varepsilon(q_1(t), q_0(t))] \\ & = \mathbb{E}_{\mathbf{h}} [\ln \varepsilon(q_1(t), q_0(t))]. \end{aligned}$$

We thus conclude that

$$\begin{aligned}\limsup_n R_n(\varepsilon_n; \mathbf{h}, \{q_{n,a}(t)\}) &= \lim_{n \rightarrow \infty} \mathbb{E}_{n,\mathbf{h}} [\ln \varepsilon_n(q_{n,1}(t), q_{n,0}(t))] \\ &\leq \mathbb{E}_{\mathbf{h}} [\ln \varepsilon(q_1(t), q_0(t))] = R(\varepsilon; \mathbf{h}, \{q_a(t)\}),\end{aligned}$$

for all $\mathbf{h} \in \mathcal{H}_1$. This proves the desired claim.

B.6. Proof of Corollary 3. By the statement of Corollary 3, $R_n(\varepsilon_n; \mathbf{h}, \{q_{n,a}(t)\}) + g(\mathbf{h}) \geq 0$ for all $\mathbf{h}, \{q_{n,a}(t)\}_n$ and n sufficiently large. Therefore,

$$\begin{aligned}\limsup_{n \rightarrow \infty} \int \{R_n(\varepsilon_n; \mathbf{h}, \{q_{n,a}(t)\}) + g(\mathbf{h})\} dw(\mathbf{h}) \\ &= \int \limsup_{n \rightarrow \infty} \{R_n(\varepsilon_n; \mathbf{h}, \{q_{n,a}(t)\}) + g(\mathbf{h})\} dw(\mathbf{h}) \\ &\leq \int \{R(\varepsilon; \mathbf{h}, \{q_a(t)\}_a) + g(\mathbf{h})\} dw(\mathbf{h}),\end{aligned}$$

for some asymptotic e-process $\varepsilon(\cdot)$, where the equality follows from the monotone convergence theorem, and the inequality follows from Theorem 4. The claim then follows by subtracting $\int g(\mathbf{h}) dw(\mathbf{h}) < \infty$ from both sides of the resulting inequality.

B.7. Proof of Corollary 4. Observe that

$$\begin{aligned}\limsup_n \mathcal{R}_n(\varepsilon_n; \{q_{n,a}(t)\}_a) \\ &\leq \inf_{\mathbf{h} \in \mathcal{H}_1} \limsup_n \left\{ \mathbb{E}_{n,\mathbf{h}_1} [\ln \varepsilon_n(q_{n,1}(t), q_{n,0}(t))] - \mathbb{E}_{n,\mathbf{h}_1} \left[\ln \frac{d\mathbb{P}_{n,\mathbf{h}_1}}{d\mathbb{P}_{n,0}}(q_{n,1}(t), q_{n,0}(t)) \right] \right\} \\ &\leq \inf_{\mathbf{h} \in \mathcal{H}_1} \left\{ \mathbb{E}_{\mathbf{h}_1} [\ln \varepsilon(q_1(t), q_0(t))] - \mathbb{E}_{\mathbf{h}_1} \left[\ln \frac{d\mathbb{P}_{\mathbf{h}_1}}{d\mathbb{P}_0}(q_1(t), q_0(t)) \right] \right\} = \mathcal{R}(\varepsilon; \{q_a(t)\}_a),\end{aligned}$$

where the second inequality follows from Theorem 4 and Assumption 5.

APPENDIX C. AUXILIARY RESULTS FOR THE PROOF OF THEOREM 1

Lemma 2. Consider the setup in Step 1 of the proof of Theorem 1. For each m, L , there exist a collection of random variables $\{q_{a,k}^{(m,L)}\}_{a,k}$ satisfying the conditions C1-C3 laid out in that step.

Proof. The construction is inductive.

Initialization

For $k = 1$, denote $\psi_{1,l} := \mathbb{E}[\tilde{\phi}_{1,l} | \{s_{a,l}\}_{a,l}]$ and

$$\varphi_{1,l} := \mathbb{E}[\tilde{\phi}_{1,l} | \{s_{1,j}\}_{j \leq l}, \{s_{0,j}\}_{j \leq \bar{c}-l+1}].$$

It is straightforward to show that the sets of random variables

$$\left\{ \tilde{\phi}_{1,l}, \{s_{1,j}\}_{j \leq l}, \{s_{0,j}\}_{j \leq \bar{c}-l+1} \right\} \text{ and } \left\{ \{s_{1,j}\}_{j > l}, \{s_{0,j}\}_{j > \bar{c}-l+1} \right\}$$

are independent of each other for any l . Indeed, this is a consequence of (A.3) and the fact that in the actual experiment

$$\{\phi_{n,1,l}, \{s_{n,1,j}\}_{j \leq l}, \{s_{n,0,j}\}_{j \leq \bar{c}-l+1}\} \text{ and } \{\{s_{n,1,j}\}_{j > l}, \{s_{n,0,j}\}_{j > \bar{c}-l+1}\}$$

are independent. Then, we have that (almost surely): (i) $\varphi_{1,l} = \psi_{1,l}$, and (ii) $\sum_{j \leq l} \varphi_{1,j} \leq \sum_{j \leq l'} \varphi_{1,j}$ for all $l' \geq l$ with $\sum_j \varphi_{1,j} = 1$. Property (i) follows from well known properties of regular conditional probabilities (see the proof of Proposition 3 in Le Cam, 1979). Property (ii) follows from (A.3) and the definition of $\varphi_{1,l}$ after noting that $\sum_{j \leq l} \phi_{n,1,j} \leq \sum_{j \leq l'} \phi_{n,1,j}$ for all $l' \geq l$ and $\sum_j \phi_{n,1,j} = 1$.

Now, take $U_1 \sim \text{Uniform}[0, 1]$ to be exogenous to $z_1(\cdot), z_0(\cdot)$, and define

$$q_{1,1}^{(m,L)} := \sum_{l=1}^{\bar{c}} \eta_l \mathbb{I} \left\{ \sum_{j \leq l-1} \varphi_{1,j} < U_1 \leq \sum_{j \leq l} \varphi_{1,j} \right\},$$

$$q_{0,1}^{(m,L)} := t_1 - q_{1,1}^{(m,L)}.$$

In the construction above, we truncate the sum at \bar{c} since $t_1 = \eta_{\bar{c}}$ and $\phi_{n,k,l} = 0$ for $l > \bar{c}$ by the definition of $q_{n,1}(\cdot)$, so $\tilde{\phi}_{1,l}, \psi_{1,l}$ and $\varphi_{1,l}$ must also be 0 almost surely for $l > \bar{c}$. From its construction, it can be verified that $\{q_{1,1}^{(m,L)} \leq \eta_l, q_{0,1}^{(m,L)} \leq \eta_{l'}\}$ is

$$\sigma \{U_1, \{s_{1,j}\}_{j \leq l+1}, \{s_{0,j}\}_{j \leq l'+1}\}$$

measurable for each l, l' such that $l + l' \geq \bar{c}$. Furthermore, the conditional law of $q_{1,1}^{(m,L)}$ given $\{s_{a,l}\}_{a,l}$ is (almost surely) equivalent to the conditional law of $\tilde{q}_{1,1}^{(m,L)}$ given $\{s_{a,l}\}_{a,l}$. This is because the conditional laws of $q_{1,1}^{(m,L)}, \tilde{q}_{1,1}^{(m,L)}$ are uniquely determined by the sets of conditional probabilities $\psi_{1,1}, \dots, \psi_{1,L}$ and $\varphi_{1,1}, \dots, \varphi_{1,L}$ which are almost surely equivalent to each other. Hence, the joint law of $(\{s_{a,l}\}_{a,l}, q_{1,1}^{(m,L)})$ is equivalent to that of $(\{s_{a,l}\}_{a,l}, \tilde{q}_{1,1}^{(m,L)})$.

Induction

Now, start with the induction hypothesis that $\{q_{a,k'}^{(m,L)}\}_{k'}$ have been constructed in a manner that satisfies conditions C1-C3 for all $k' < k$. We show how to extend the construction to $k' = k$ so that conditions C1-C3 continue to be satisfied.

It is useful to note that if $\tilde{q}_{1,k}^{(m,L)} = \eta_l$, then $\tilde{q}_{0,k}^{(m,L)} = t_k - \eta_l = \eta_{k\bar{c}-l}$. Denote

$$\begin{aligned}\Phi_n^{(k,l)} &\equiv \left\{ \phi_{n,\tilde{k},\tilde{l}} : \tilde{l} \leq l, \tilde{k}\bar{c} - \tilde{l} \leq k\bar{c} - l, \tilde{k} \leq k-1 \right\}, \\ \tilde{\Phi}^{(k,l)} &\equiv \left\{ \tilde{\phi}_{\tilde{k},\tilde{l}} : \tilde{l} \leq l, \tilde{k}\bar{c} - \tilde{l} \leq k\bar{c} - l, \tilde{k} \leq k-1 \right\}.\end{aligned}$$

Intuitively, $\tilde{\Phi}_n^{(k,l)}$ represents the collection of $\phi_{n,\tilde{k},\tilde{l}}$ variables that require less ‘information’ to determine than $\phi_{n,k,l}$. Define the σ -algebras

$$\begin{aligned}\mathcal{L}_{k,l} &\equiv \sigma \left\{ \{s_{1,j}\}_j, \{s_{0,j}\}_j, \left\{ \tilde{q}_{a,\tilde{k}}^{(m,L)} : \tilde{k} \leq k-1 \right\}_a \right\}, \\ \mathcal{M}_{k,l} &\equiv \sigma \left\{ \{s_{1,j}\}_j, \{s_{0,j}\}_j, \tilde{\Phi}^{(k,l)} \right\}, \\ \mathcal{H}_{k,l} &\equiv \sigma \left\{ \{s_{1,j}\}_{j \leq l}, \{s_{0,j}\}_{j \leq k\bar{c}-l+1}, \tilde{\Phi}^{(k,l)} \right\},\end{aligned}$$

and let $B_{n,k,l}, \tilde{B}_{k,l}, B_{k,l}$ denote the events

$$\begin{aligned}\tilde{B}_{n,k,l} &\equiv \left\{ \tilde{q}_{n,1,\tilde{k}}^{(m,L)} \leq \eta_l, \tilde{q}_{n,0,\tilde{k}}^{(m,L)} \leq \eta_{k\bar{c}-l}, \forall \tilde{k} \leq k-1 \right\}, \\ \tilde{B}_{k,l} &\equiv \left\{ \tilde{q}_{1,\tilde{k}}^{(m,L)} \leq \eta_l, \tilde{q}_{0,\tilde{k}}^{(m,L)} \leq \eta_{k\bar{c}-l}, \forall \tilde{k} \leq k-1 \right\}, \\ B_{k,l} &\equiv \left\{ q_{1,\tilde{k}}^{(m,L)} \leq \eta_l, q_{0,\tilde{k}}^{(m,L)} \leq \eta_{k\bar{c}-l}, \forall \tilde{k} \leq k-1 \right\}.\end{aligned}$$

Also, take $\mathcal{L}_{k,l}^+$ and $\mathcal{L}_{k,l}^-$ to be the σ -algebras corresponding to the restriction of $\mathcal{L}_{k,l}$ to $\tilde{B}_{k,l}$ and $\tilde{B}_{k,l}^c$, respectively. The quantities $\mathcal{M}_{k,l}^+, \mathcal{M}_{k,l}^-$ and $\mathcal{H}_{k,l}^+, \mathcal{H}_{k,l}^-$ are defined analogously. Observe that $\mathcal{L}_{k,l}^+ \equiv \mathcal{M}_{k,l}^+$. This is because, when $\tilde{B}_{k,l}$ holds, all the $\{\tilde{\phi}_{\tilde{k},l}\}_{\tilde{k} \leq k-1}$ random variables outside the collection $\tilde{\Phi}^{(k,l)}$ necessarily take on the value 0, so they do not provide any additional information.

Define $\vartheta_{k,l} := \mathbb{E} [\tilde{\phi}_{k,l} | \mathcal{L}_{k,l}]$ and observe that

$$\begin{aligned}\vartheta_{k,l} &= \mathbb{I}\{\tilde{B}_{k,l}\} \mathbb{E} [\tilde{\phi}_{k,l} | \mathcal{L}_{k,l}^+] + \mathbb{I}\{\tilde{B}_{k,l}^c\} \mathbb{E} [\tilde{\phi}_{k,l} | \mathcal{L}_{k,l}^-] \text{ a.s.} \\ &= \mathbb{I}\{\tilde{B}_{k,l}\} \mathbb{E} [\tilde{\phi}_{k,l} | \mathcal{M}_{k,l}^+] \text{ a.s.}\end{aligned}$$

The second step uses the fact $\tilde{q}_{a,k}^{(m,L)}$ is almost surely non-decreasing in k as it is the weak limit of $q_{n,a,k}^{(m,L)}$, which is non-decreasing in k . Hence, conditional on $\tilde{B}_{k,l}^c$, we have $\tilde{\phi}_{k,l} = 0$ a.s., implying $\mathbb{E}[\tilde{\phi}_{k,l} | \mathcal{L}_{k,l}^-] = 0$ a.s.

Set $\psi_{k,l} := \mathbb{E}[\tilde{\phi}_{k,l} | \mathcal{M}_{k,l}^+]$ and $\varphi_{k,l} := \mathbb{E}[\tilde{\phi}_{k,l} | \mathcal{H}_{k,l}^+]$. Now, in the actual experiment, the collection of random variables

$$\left\{ \phi_{n,k,l}, \Phi_n^{(k,l)}, \{s_{n,1,j}\}_{j \leq l}, \{s_{n,0,j}\}_{j \leq k\bar{c}-l+1}, \mathbb{I}\{B_{n,k,l}\} \right\}$$

and $\{\{s_{n,1,j}\}_{j > l}, \{s_{n,0,j}\}_{j > k\bar{c}-l+1}\}$

are independent of each other. Combined with (A.3) and the properties of weak convergence, we conclude

$$\left\{ \tilde{\phi}_{k,l}, \tilde{\Phi}^{(k,l)}, \{s_{1,j}\}_{j \leq l}, \{s_{0,j}\}_{j \leq k\bar{c}-l+1}, \mathbb{I}\{\tilde{B}_{k,l}\} \right\}$$

and $\{\{s_{1,j}\}_{j > l}, \{s_{0,j}\}_{j > k\bar{c}-l+1}\}$

are also independent of each other for any l . Hence, by similar arguments as in the initialization step, it follows that almost surely: (i) $\varphi_{k,l} = \psi_{k,l}$, and (ii) $\sum_{j \leq l} \varphi_{k,j} \leq \sum_{j \leq l'} \varphi_{k,j}$ for all $l' \geq l$ with $\sum_j \varphi_{k,j} = 1$. The Doob-Dynkin theorem states that we may take $\varphi_{k,l}$ to be a measurable function of

$$\{s_{1,j}\}_{j \leq l}, \{s_{0,j}\}_{j \leq k\bar{c}-l+1}, \tilde{\Phi}^{(k,l)},$$

the random variables generating $\mathcal{H}_{k,l}$. Denote this function by $\varphi_{k,l}(\cdot)$.

We define $q_{a,k}^{(m,L)}$ on a new probability space (i.e., separate from the space in which $\tilde{q}_{a,l}^{(m,L)}$ reside) containing $\{z_1(\cdot), z_0(\cdot), U_1, \dots, U_{2^m}\}$, where U_1, \dots, U_{2^m} are iid Uniform[0, 1] and independent of $z_1(\cdot), z_0(\cdot)$. Formally, given some values for $\{q_{1,\tilde{k}}^{(m,L)}\}_{\tilde{k} \leq k-1}$ —themselves functions of $z_1(\cdot), z_0(\cdot), U_1, \dots, U_{k-1}$ by the induction hypothesis—we set:

$$q_{1,k}^{(m,L)} = \sum_{l=0}^{k\bar{c}} \eta_l \mathbb{I}\{B_{k,l}\} \cdot \mathbb{I} \left\{ \sum_{j \leq l-1} \varphi_{k,j}(\cdot) < U_k \leq \sum_{j \leq l} \varphi_{k,j}(\cdot) \right\},$$

$$q_{0,k}^{(m,L)} = t_k - q_{1,k}^{(m,L)}.$$

In the equation above, the functions $\{\varphi_{k,l}(\cdot)\}_l$ now take as inputs the corresponding quantities on the new probability space, i.e., $\varphi_{k,l}(\cdot)$ is applied on

$$\{s_{1,j}\}_{j \leq l}, \{s_{0,j}\}_{j \leq k\bar{c}-l+1}, \Phi^{(k,l)}, \quad (\text{C.1})$$

where

$$\begin{aligned} \Phi^{(k,l)} &:= \left\{ \phi_{\tilde{k},\tilde{l}} : \tilde{l} \leq l, \tilde{k}\bar{c} - \tilde{l} \leq k\bar{c} - l, \tilde{k} \leq k - 1 \right\}, \text{ and} \\ \phi_{\tilde{k},\tilde{l}} &:= \mathbb{I} \left\{ q_{1,\tilde{k}}^{(m,L)} = \eta_{\tilde{l}} \right\}. \end{aligned} \quad (\text{C.2})$$

We verify below that $q_{a,k}^{(m,L)}$, so defined, satisfies conditions C1-C3.

Condition C1 is satisfied by construction since $q_{1,k}^{(m,L)} + q_{0,k}^{(m,L)} = t_k$.

Next, we verify Condition C2. Consider the event $\{q_{1,k}^{(m,L)} \leq \eta_l, q_{0,k}^{(m,L)} \leq \eta_{l'}\}$ for any $l + l' \geq k\bar{c}$. Because $q_{1,k}^{(m,L)} + q_{0,k}^{(m,L)} = t_k$, the definition of $q_{a,k}^{(m,L)}$ above implies

$$\begin{aligned} \{q_{1,k}^{(m,L)} \leq \eta_l, q_{0,k}^{(m,L)} \leq \eta_{l'}\} &\equiv \left\{ \eta_{k\bar{c}+l-l'} \leq q_{1,k}^{(m,L)} \leq \eta_l \right\} \\ &\equiv \bigcup_{i=k\bar{c}+l-l'}^l \left\{ q_{1,k}^{(m,L)} = \eta_i \right\} \\ &\equiv \bigcup_{i=k\bar{c}+l-l'}^l \left\{ \sum_{j \leq i-1} \varphi_{k,j}(\cdot) < U_k \leq \sum_{j \leq i} \varphi_{k,j}(\cdot) \right\} \cap B_{k,l}. \end{aligned} \quad (\text{C.3})$$

Since $\{q_{a,\tilde{k}}^{(m,L)}\}_{\tilde{k} \leq k-1}$ satisfy Condition C2 (by the induction hypothesis),

$$B_{k,l} \in \sigma \{U_1, \dots, U_k, \{s_{1,j}\}_{j \leq l+1}, \{s_{0,j}\}_{j \leq l'+1}\}$$

for each $l + l' \geq k\bar{c}$. Furthermore, the set of $\varphi_{k,l}(\cdot)$ functions used in (C.3) are from the collection $\{\varphi_{k,j}(\cdot)\}_{j=k\bar{c}+l-l'-1}^l$. By (C.1), these are all measurable functions of the maximal set

$$\{s_{1,j}\}_{j \leq l}, \{s_{0,j}\}_{j \leq k\bar{c}-l+1}, \left\{ \phi_{\tilde{k},\tilde{l}} : \tilde{l} \leq l, \tilde{k}\bar{c} - \tilde{l} \leq l' + 1, \tilde{k} \leq k - 1 \right\}. \quad (\text{C.4})$$

From the definition of $\phi_{k,l}$ (C.2) and the induction hypothesis, some straightforward algebra reveals that the above random variables are all, in turn,

$$\sigma \{U_1, \dots, U_k, \{s_{1,j}\}_{j \leq l+1}, \{s_{0,j}\}_{j \leq l'+1}\}$$

measurable as well. This completes the verification of Condition C2.

It remains to verify Condition C3. Note that U_k is independent of $\{s_{a,l}\}_{a,l}$. Hence, for each η_l , the construction of $q_{1,k}^{(m,L)}$ implies

$$\begin{aligned} & \mathbb{P}\left(q_{1,k}^{(m,L)} = \eta_l \mid \{s_{a,j}\}_{a,j}, \left\{\tilde{q}_{a,k'}^{(m,L)} : k' \leq k-1\right\}_a, B_{k,l}\right) \\ &= \mathbb{P}\left(q_{1,k}^{(m,L)} = \eta_l \mid \{s_{a,j}\}_{a,j}, \Phi^{(k,l)}, B_{k,l}\right) \\ &= \mathbb{P}\left(q_{1,k}^{(m,L)} = \eta_l \mid \{s_{1,j}\}_{j \leq l}, \{s_{0,j}\}_{j \leq k-\bar{l}+1}, \Phi^{(k,l)}, B_{k,l}\right) \\ &= \varphi_{k,l}, \end{aligned}$$

which is the same (a.s.) as $\psi_{k,l}$, the conditional probability of $\{\tilde{q}_{1,k}^{(m,L)} = \eta_l\}$ given $\mathcal{L}_{k,l}^+$. At the same time, $\{q_{1,k}^{(m,L)} = \eta_l\}$ never occurs when $B_{k,l}^c$ occurs, which matches the fact that the conditional probability of $\{\tilde{q}_{1,k}^{(m,L)} = \eta_l\}$ given $\tilde{B}_{k,l}^c$ is 0. So, overall, we conclude that the conditional law of $q_{1,k}^{(m,L)}$ given $\{s_{a,l}\}_{a,l}$ and $\{\tilde{q}_{a,k'}^{(m,L)} : k' \leq k\}_a$ is the same as the conditional law of $\tilde{q}_{a,k}^{(m,L)}$ given $\mathcal{L}_{k,l}$. Combined with the induction hypothesis, it then follows that the joint law of $(\{s_{a,l}\}_{a,l}, q_{1,1}^{(m,L)}, \dots, q_{1,k}^{(m,L)})$ is equivalent to that of $(\{s_{a,l}\}_{a,l}, \tilde{q}_{1,1}^{(m,L)}, \dots, \tilde{q}_{1,k}^{(m,L)})$. \square

APPENDIX D. EQUIVALENCE OF OUT-OF-SAMPLE REGRET

In contrast to in-sample regret, out-of-sample (or simple) regret measures the expected difference between the welfare from a chosen action and that of the optimal action, evaluated on new, unseen data. Specifically, suppose that at the end of the adaptive experiment, the DM is tasked with specifying a treatment decision $\boldsymbol{\delta}_n \equiv (\delta_{n,1}, \delta_{n,0}) \in \mathcal{S}^2$ to be applied on the entire population. Here \mathcal{S}^2 denotes the 2-dimensional simplex. The out-of-sample frequentist regret of this decision is then defined as

$$W_n^o(\mathbf{h}) = \sqrt{n} \left\{ \max_a \mu_{n,a}(\mathbf{h}) - \sum_a \mu_{n,a}(\mathbf{h}) \mathbb{E}_{n,\mathbf{h}}[\delta_{n,a}] \right\}.$$

Clearly, $\boldsymbol{\delta}_n$ must be $\mathcal{I}_{n,1} := \mathcal{G}_{n,q_{n,1}(1),q_{n,0}(1)}$ measurable.

Analogously, in the limit experiment, the out-of-sample frequentist regret is defined as

$$W^o(\mathbf{h}) = \max_a \dot{\mu}_a^\top \mathbf{h} - \sum_a \dot{\mu}_a^\top \mathbf{h} \mathbb{E}_{\mathbf{h}}[\delta_a],$$

where $\boldsymbol{\delta} \equiv (\delta_1, \delta_0) \in \mathcal{S}^2$ is \mathcal{I}_1 -measurable, and $\dot{\mu}_a(\cdot)$ is defined in Section 5. We then have the following analogue to Theorem 3.

Theorem 5. Suppose Assumptions 1, 2 and 4 hold. Let $\{x_a(\cdot), q_a(\cdot)\}_a$ denote the weak limit of $\{x_{n,a}(\cdot), q_{n,a}(\cdot)\}_a$ under a sequence of policy rules $\{\pi_{n,j}\}_j$ in the actual experiment, and $\{\delta_n\}_n$ a sequence of treatment decisions with corresponding out-of-sample regret $W_n^o(\mathbf{h})$. Then, there exists a subsequence $\{\delta_{n_k}\}_k$ and a treatment decision $\boldsymbol{\delta}$ in the limit experiment, depending only on $\{x_a(1), q_a(1)\}_a$, with out-of-sample regret $W^o(\mathbf{h})$, such that $\lim_{k \rightarrow \infty} W_{n_k}^o(\mathbf{h}) = W^o(\mathbf{h})$ for each \mathbf{h} .

Proof. As $\{\delta_n\}_n$ is uniformly bounded, it is tight. Combined with (3.1), Assumption 2 and Theorem 1, it follows that the joint

$$(\delta_n, \hat{\varphi}(\mathbf{h}; q_{n,1}(1), q_{n,0}(1)))$$

is also tight. Hence, by Prohorov's theorem, given any sequence $\{n\}$, there exists a further sub-sequence $\{n_k\}$ —represented as $\{n\}$ for ease of notation—such that

$$\begin{aligned} \left(\begin{array}{c} \delta_n \\ \hat{\varphi}(\mathbf{h}; q_{n,1}(1), q_{n,0}(1)) \end{array} \right) &\xrightarrow[\mathbb{P}_{n,0}]{} \left(\begin{array}{c} \bar{\boldsymbol{\delta}} \\ V \end{array} \right); \text{ where} \\ V &\sim \exp \sum_a \left\{ h^{(a)\top} I_a^{1/2} x_a(1) - \frac{q_a(1)}{2} h^{(a)\top} I_a h^{(a)} \right\}, \end{aligned} \quad (\text{D.1})$$

and $\bar{\boldsymbol{\delta}} \in [0, 1]$ is some tight limit of δ_n . Therefore, by similar arguments as in the proof of Theorem 2,

$$\delta_n \xrightarrow[\mathbb{P}_{n,\mathbf{h}}]{} \mathcal{L}; \text{ where } \mathcal{L}(B) := E \left[\mathbb{I}\{\bar{\boldsymbol{\delta}} \in B\} V \right] \forall B \in \mathcal{B}(\mathbb{R}^2). \quad (\text{D.2})$$

Define $\boldsymbol{\delta} = E \left[\bar{\boldsymbol{\delta}} | \{x_a(1), q_a(1)\}_a \right]$. By construction, $\boldsymbol{\delta}$ is a valid treatment policy in the limit experiment, and it is also \mathcal{I}_1 -measurable. Furthermore, by (D.2),

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}_{n,\mathbf{h}}[\delta_{n,a}] &= E \left[\bar{\delta}_a e^{\sum_a \left\{ h^{(a)\top} I_a^{1/2} x_a(1) - \frac{q_a(1)}{2} h^{(a)\top} I_a h^{(a)} \right\}} \right] \\ &= E \left[\delta_a e^{\sum_a \left\{ h^{(a)\top} I_a^{1/2} x_a(1) - \frac{q_a(1)}{2} h^{(a)\top} I_a h^{(a)} \right\}} \right] = \mathbb{E}_{\mathbf{h}}[\delta_a] \forall a, \end{aligned}$$

where the second equality follows by the law of iterated expectations, and the last equality follows by the Girsanov theorem.

We have thereby shown that $\mathbb{E}_{n,\mathbf{h}}[\delta_{n,a}] \rightarrow \mathbb{E}_{\mathbf{h}}[\delta_a]$ for each \mathbf{h}, a . Combined with Assumption 4, this implies $\lim_{n \rightarrow \infty} W_n^o(\mathbf{h}) = W^o(\mathbf{h})$ for each \mathbf{h} , which proves the desired claim. \square

Theorem 5 is stronger than Theorem 3 in that it reduces the set of sufficient statistics for treatment assignment to $\{x_a(1), q_a(1)\}_a$, i.e., only the terminal values of the score and allocation processes are relevant.

APPENDIX E. ADDITIONAL RESULTS ON ANYTIME-VALID INFERENCE

E.1. Verifying the conditions of Corollary 3 for $\varepsilon_n^*(\cdot)$ from (6.4). We show that the asymptotically mGO optimal e-process $\varepsilon_n^*(\cdot)$ satisfies the requirements for Corollary 3 under the following primitive conditions:

- Assumption 6.** (i) There exists $p > 0$ independent of n, \mathbf{h} such that $\mathbb{E}_{n,\mathbf{h}} \left[\left| \psi(Y_i^{(a)}) \right|^{2+p} \right] < \infty$ for each a, \mathbf{h} .
(ii) For each a , $\sqrt{n} \mathbb{E}_{n,\mathbf{h}} \left[\psi(Y_i^{(a)}) \right] = I_a h + \delta_n |h^{(a)}|$, where $\delta_n \rightarrow 0$ is independent of \mathbf{h} .
(iii) The weighting function $w(\cdot)$ satisfies $\int e^{c\mathbf{h}^\top \mathbf{h}} dw(\mathbf{h}) \leq M < \infty$ for some $c > 0$.

Assumption 6(i) is a mild regularity condition. Assumption 6(ii) follows from quadratic mean differentiability (Assumption 1). Assumption 6(iii) requires the weighting function to have sub-Gaussian tails. This is natural since $\varepsilon_n^*(\cdot)$ is based on integrating an exponential term with respect to $w(\mathbf{h})$.

We verify the various requirements for Corollary 3 below:

Weak convergence. Based on the form of $\varepsilon_n^*(\cdot)$, Theorem 1 and standard weak convergence arguments imply $\{z_{n,a}(\cdot), \varepsilon_n(\cdot, \cdot)\}_a$ converges weakly under $\mathbb{P}_{n,0}$. This verifies the first part of Assumption 4.

Uniform Integrability. We now show that $\sup_{q_1, q_0} \ln \varepsilon_n^*(q_1, q_0)$ is uniformly integrable under $\mathbb{P}_{n,\mathbf{h}}$, which verifies the second part of Assumption 4.

Denote $\tilde{\mathbf{z}}_n(\mathbf{q}) = [z_{n,1}(q_1)^\top I_1^{1/2}, z_{n,0}(q_0)^\top I_0^{1/2}]^\top$, $\mathbf{I} = \text{diag}(I_1, I_0)$ and $\mathbf{I}_q = \text{diag}(q_1 I_1, q_0 I_0)$.¹³ Then,

$$\begin{aligned} \varepsilon_n^*(q_1, q_0) &= \int e^{\mathbf{h}^\top \tilde{\mathbf{z}}_n(\mathbf{q}) - \frac{1}{2} \mathbf{h}^\top \mathbf{I}_q \mathbf{h}} dw(\mathbf{h}) \\ &= \int e^{\mathbf{h}^\top \tilde{\mathbf{z}}_n(\mathbf{q}) - \frac{1}{2} \mathbf{h}^\top (\mathbf{I}_q + 2c \cdot \text{Id}) \mathbf{h}} \cdot e^{c\mathbf{h}^\top \mathbf{h}} dw(\mathbf{h}), \end{aligned}$$

¹³Here, $\text{diag}(A, B)$ denotes the block diagonal matrix with diagonal (matrix) elements A, B .

where Id denotes the identity matrix. Define $\boldsymbol{\Lambda} = \mathbf{I}_q + 2c \cdot \text{Id}$, and note that by completing the square,

$$\begin{aligned}\varepsilon_n^*(q_1, q_0) &= e^{\frac{1}{2}\tilde{\mathbf{z}}_n(\mathbf{q})^\top \boldsymbol{\Lambda}^{-1}\tilde{\mathbf{z}}_n(\mathbf{q})} \int e^{-\frac{1}{2}|\boldsymbol{\Lambda}^{1/2}\mathbf{h} - \boldsymbol{\Lambda}^{-1/2}\tilde{\mathbf{z}}_n(\mathbf{q})|^2} e^{c\mathbf{h}^\top \mathbf{h}} dw(\mathbf{h}) \\ &\leq e^{\frac{1}{2}\tilde{\mathbf{z}}_n(\mathbf{q})^\top \boldsymbol{\Lambda}^{-1}\tilde{\mathbf{z}}_n(\mathbf{q})} \int e^{c\mathbf{h}^\top \mathbf{h}} dw(\mathbf{h}) \\ &\leq M e^{\frac{1}{2}\tilde{\mathbf{z}}_n(\mathbf{q})^\top \boldsymbol{\Lambda}^{-1}\tilde{\mathbf{z}}_n(\mathbf{q})} \leq M e^{\frac{1}{4c}\tilde{\mathbf{z}}_n(\mathbf{q})^\top \tilde{\mathbf{z}}_n(\mathbf{q})}\end{aligned}$$

where the third step follows by Assumption 6(iii), and the last step makes use of the fact that \mathbf{I}_q is positive semi-definite for any q_1, q_0 .

Hence, $\sup_{q_1, q_0} \ln \varepsilon_n^*(q_1, q_0)$ is uniformly integrable under $\mathbb{P}_{n, \mathbf{h}}$ as long as

$$\sup_{q_1, q_0} \tilde{\mathbf{z}}_n(\mathbf{q})^\top \tilde{\mathbf{z}}_n(\mathbf{q}) = \sup_{q_1} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nq_1 \rfloor} \psi(Y_i^{(1)}) \right|^2 + \sup_{q_0} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nq_0 \rfloor} \psi(Y_i^{(0)}) \right|^2$$

is uniformly integrable under $\mathbb{P}_{n, \mathbf{h}}$. It therefore suffices to show that each term, $\sup_{q_a} \left| n^{-1/2} \sum_{i=1}^{\lfloor nq_a \rfloor} \psi(Y_i^{(a)}) \right|^2$, is uniformly integrable. Take $a = 1$ without loss of generality and note that by Assumption 6(ii),

$$\begin{aligned}&\sup_{q_1 \in [0, 1]} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nq_1 \rfloor} \psi(Y_i^{(1)}) \right|^2 \\ &\leq 2 \sup_{q_1 \in [0, 1]} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nq_1 \rfloor} \left\{ \psi(Y_i^{(1)}) - \mathbb{E}_{n, \mathbf{h}} [\psi(Y_i^{(1)})] \right\} \right|^2 + 4 \left| I_a h^{(a)} \right|^2 + 4\delta_n |h^{(a)}|^2.\end{aligned}\quad (\text{E.1})$$

Define $A_{n,i} := \psi(Y_i^{(1)}) - \mathbb{E}_{n, \mathbf{h}} [\psi(Y_i^{(1)})]$ and observe that $M_k := n^{-1/2} \sum_{i=1}^k A_{n,i}$ is a martingale under $\mathbb{P}_{n, \mathbf{h}}$. Then, for any $p > 0$,

$$\begin{aligned}&\mathbb{E}_{n, \mathbf{h}} \left[\sup_{q_1 \in [0, 1]} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor nq_1 \rfloor} A_{n,i} \right|^{2+p} \right] = \mathbb{E}_{n, \mathbf{h}} \left[\sup_{k \leq n} |M_k|^{2+p} \right] \\ &\leq \left(\frac{2+p}{1+p} \right)^{2+p} \mathbb{E}_{n, \mathbf{h}} [|M_n|^{2+p}] \lesssim \mathbb{E}_{n, \mathbf{h}} [|A_{n,i}|^{2+p}] < \infty,\end{aligned}$$

where the second step follows by Doob's maximal inequality, the penultimate step uses the Marcinkiewicz–Zygmund inequality, and the last step follows from Assumption 6(i). This proves that $\sup_{q_1 \in [0, 1]} \left| n^{-1/2} \sum_{i=1}^{\lfloor nq_1 \rfloor} A_{n,i} \right|^2$ is uniformly integrable, and therefore, in view of (E.1), that $\sup_{q_1} \left| n^{-1/2} \sum_{i=1}^{\lfloor nq_1 \rfloor} \psi(Y_i^{(1)}) \right|^2$ is uniformly integrable as well.

Lower bound on GRO value. We now show that the GRO value, $R_n(\varepsilon_n^*; \mathbf{h}, \{q_{n,a}(t)\}_a)$, of $\varepsilon_n^*(\cdot)$ is always non-negative for n large enough. Thus, the function $g(\cdot)$ in the statement of Corollary 3 can be set to 0.

To prove this, observe that by Jensen's inequality,

$$\ln \varepsilon_n^*(q_{n,1}(t), q_{n,0}(t)) \geq \int \sum_a \left\{ h^{(a)\top} I_a^{1/2} z_{n,a}(q_{n,a}) - \frac{q_{n,a}}{2} h^{(a)\top} I_a h^{(a)} \right\} dw(\mathbf{h}).$$

Consequently, Fubini's theorem implies

$$R_n(\varepsilon_n^*; \mathbf{h}, \{q_{n,a}(t)\}_a) \geq \int \sum_a \left\{ h^{(a)\top} I_a^{1/2} \mathbb{E}_{n,\mathbf{h}}[z_{n,a}(q_{n,a})] - \frac{\mathbb{E}_{n,\mathbf{h}}[q_{n,a}]}{2} h^{(a)\top} I_a h^{(a)} \right\} dw(\mathbf{h}).$$

Now, Wald's identity and Assumption 6(ii) imply

$$\mathbb{E}_{n,\mathbf{h}}[z_{n,a}(q_a)] = I_a^{-1/2} \frac{n \mathbb{E}_{n,\mathbf{h}}[q_{n,a}] \mathbb{E}_{n,\mathbf{h}}[\psi(Y_i^{(a)})]}{\sqrt{n}} = \mathbb{E}_{n,\mathbf{h}}[q_{n,a}] \left(I_a^{1/2} h^{(a)} + \epsilon_n |h^{(a)}| \right),$$

where $\epsilon_n \rightarrow 0$. Since $h^{(a)\top} I_a h^{(a)} \geq 2\epsilon_n |h^{(a)}|^2 \forall a$ for n large enough, we thus have

$$R_n(\varepsilon_n^*; \mathbf{h}, \{q_{n,a}(t)\}_a) \geq \sum_a \int \frac{\mathbb{E}_{n,\mathbf{h}}[q_{n,a}]}{2} \left(h^{(a)\top} I_a h^{(a)} - 2\epsilon_n |h^{(a)}|^2 \right) dw(\mathbf{h}) \geq 0.$$

E.2. Locally REGROW optimal e-processes. In this section, we construct e-processes that are locally optimal against fixed values of $\{q_a\}_a$. These e-processes therefore achieve REGROW optimality with respect to non-adaptive sampling designs where the allocations are fixed in advance.

For a given pair (q_1, q_0) , equation (6.6) implies that the optimal weighting function $w_{(q_1, q_0)}^*$ is obtained by solving

$$w_{(q_1, q_0)}^* = \arg \max_{w \in \Delta(\mathcal{H}_1)} \text{KL}_{(q_1, q_0)}(p_{\mathbf{h}} \cdot w \parallel p_w \cdot w) = \arg \max_{w \in \Delta(\mathcal{H}_1)} I_{(q_1, q_0)}(w; p_w), \quad (\text{E.2})$$

where $\mathbb{P}_{\mathbf{h}}^{(q_1, q_0)}$ corresponds to an independent bivariate-normal distribution

$$\prod_a \mathcal{N}(I_a^{1/2} q_a h^{(a)}, q_a).$$

The channel capacity and w^* depend on the structure of the compact set \mathcal{H}_1 .

As a leading example—following Grünwald et al. (2024)—consider the case where each h_a is scalar and

$$\mathcal{H}_1 \equiv \left\{ \mathbf{h} : |h^{(a)}| \leq K \ \forall a \right\},$$

for some constant $K < \infty$. In this setting, the optimization problem (E.2) factorizes across arms, yielding a product form solution $w^* = \prod_a w_a^*(h^{(a)})$, and the overall channel capacity becomes the sum of individual channel capacities for each arm. Classical results from Smith (1971) establish that the least-favorable distribution $w_a^*(\cdot)$ is always discrete. For low signal to noise ratios, specifically, when $K/I_a^{1/2}q_a \lesssim 1.4$, $w_a^*(\cdot)$ reduces to a symmetric two-point prior supported on $\{-K, K\}$. However, when q_a falls below the critical threshold $K/(1.4I_a^{1/2})$, the support of w_a^* expands to include more than two points and must be computed via a finite-dimensional convex program, as described in Smith (1971).

Importantly, because w^* depends on (q_1, q_0) , there is no e-process that is simultaneously REGROW optimal for all possible allocation pairs. Nevertheless, when (q_1, q_0) are sufficiently large, w^* stabilizes at the symmetric two-point prior on $\{-K, K\}$. The corresponding e-process

$$\varepsilon^*(q_1, q_0) = \prod_a \frac{1}{2} \sum_{h=\pm K} \exp \sum_a \left\{ h I_a^{1/2} z_a(q_a) - \frac{q_a}{2I_a} h^2 \right\}$$

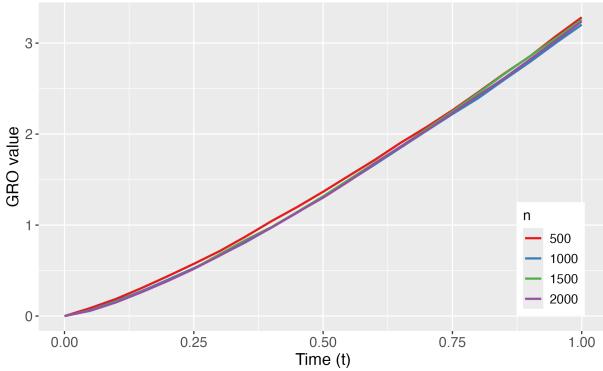
is therefore REGROW optimal at these values. As before, a corresponding finite-sample approximation can be constructed by replacing $z_a(\cdot)$ with its empirical counterpart $z_{n,a}(\cdot)$.

E.3. Testing parameters corresponding to individual arms. Mirroring the setup of Section 6.6, suppose we are interested in conducting an anytime-valid test of the null hypothesis $H_0 : \theta^{(1)} = \bar{\theta}$ against the two-sided alternative $H_1 : \theta^{(1)} \neq \bar{\theta}$. Since $\theta^{(0)}$ is unrestricted, this corresponds to a testing problem with a composite null $\Theta_0 \equiv \{(\bar{\theta}, \theta^{(0)}) : \theta^{(0)} \in \mathbb{R}\}$ and a composite alternative $\Theta_1 \equiv \{(\theta^{(1)}, \theta^{(0)}) : \theta^{(1)} \neq \bar{\theta}, \theta^{(0)} \in \mathbb{R}\}$.

We take the reference parameter vector to be $\boldsymbol{\theta}_0 = (\bar{\theta}, \bar{\theta})$, and consider local alternatives of the form $\boldsymbol{\theta} = \boldsymbol{\theta}_0 + \mathbf{h}/\sqrt{n}$. The partitions Θ_0, Θ_1 of the space of $\boldsymbol{\theta}$ induce a corresponding partition of the local parameter space which we denote by

$$\mathcal{H}_0 \equiv \{(0, h^{(0)}) : h^{(0)} \in \mathbb{R}\} \quad \text{and} \quad \mathcal{H}_1 \equiv \{(h^{(1)}, h^{(0)}) : h^{(1)} \neq 0, h^{(0)} \in \mathbb{R}\}.$$

Assume a prior (or weight function) $w_1(h^{(1)})$ is placed over the alternative values of $h^{(1)}$. Fix any $h^{(0)} \in \mathbb{R}$. Then, by (6.4), the unique mGRO-optimal e-process



Note: The figure displays the GRO value as a function of time for the e-process in (6.7) under Thompson Sampling, at the local alternative $(h^{(1)}, h^{(0)}) = (1/\sqrt{n}, 0)$.

FIGURE E.1. Evolution of GRO under Thompson Sampling

for testing the simple null $\bar{H}_0 : \mathbf{h} = (0, h^{(0)})$ against the composite alternative $\bar{H}_1 : \mathbf{h} \in \{(h^{(1)}, h^{(0)}) : h^{(1)} \neq 0\}$, given $w_1(\cdot)$, is independent of observations from arm 0 and takes the form:

$$\varepsilon^*(q_1(t)) = \int \exp \left\{ h^{(1)\top} I_1^{1/2} z_1(q_1(t)) - \frac{q_1(t)}{2} h^{(1)\top} I_1 h^{(1)} \right\} dw_1(h^{(1)}).$$

This process is clearly a valid e-process for testing the composite null $\mathbf{h} \in \mathcal{H}_0$ against the composite alternative $\mathbf{h} \in \mathcal{H}_1$. Moreover, since $\varepsilon^*(q_1(t))$ is mGRO-optimal for each fixed value of $h^{(0)}$, it follows that it is also mGRO-optimal for testing $\mathbf{h} \in \mathcal{H}_0$ against $\mathbf{h} \in \mathcal{H}_1$, when there is a weight function $w_1(\cdot)$ over $h^{(1)}$.

E.4. Additional simulation results. Figure E.1 replicates the analysis in Panel B of Figure 6.1, but replaces the UCB algorithm with Thompson Sampling (TS). The resulting curves again exhibit striking stability across values of n . Notably, the e-process defined in equation (6.7) achieves uniformly higher GRO values under TS than under UCB. This difference arises from the fact that we only consider alternatives where $\theta^{(1)} = 0.1 + 1/\sqrt{n}$ and $\theta^{(0)} = 0.1$. Under this class of alternatives, UCB allocates fewer observations, on average, to arm 1—the arm being tested—than TS does. As a result, the TS-based allocation yields a more powerful e-process in this setting.