Optimal Conditional Inference in Adaptive Experiments

Jiafeng Chen[1] and Isaiah Andrews[2]

Abstract. We study batched bandit experiments and consider the problem of inference conditional on the realized stopping time, assignment probabilities, and target parameter, where all of these may be chosen adaptively using information up to the last batch of the experiment. Absent further restrictions on the experiment, we show that inference using only the results of the last batch is optimal. When the adaptive aspects of the experiment are known to be location-invariant, in the sense that they are unchanged when we shift all batch-arm means by a constant, we show that there is additional information in the data, captured by one additional linear function of the batch-arm means. In the more restrictive case where the stopping time, assignment probabilities, and target parameter are known to depend on the data only through a collection of polyhedral events, we derive computationally tractable and optimal conditional inference procedures.

Keywords. Conditional inference, sufficient statistic, bandit, selective inference, uniform inference

[1]Department of Economics, Harvard University and Harvard Business School.

[2]Department of Economics, Massachusetts Institute of Technology.

E-mail addresses: `jiafengchen@g.harvard.edu; iandrews@mit.edu`.

Date: September 22, 2023. An extended abstract of this paper was presented at CODE@MIT 2021. We thank Kei Hirano, Lucas Janson, Lihua Lei, Jonathan Roth, Brad Ross, David Ritzwoller, Pedro Sant'Anna, and participants at the Microsoft Research causal inference reading group and at CODE@MIT for comments.

1. Introduction

Consider a batched bandit experiment (Zhang, Janson and Murphy, 2020; Hirano and Porter, 2023), where at each batch $t = 1, 2, \ldots, T_0$ an experimenter either continues or ends the experiment. If they continue the experiment, n_t units are independently randomized among K different arms, where the probability that a unit is assigned to arm k in batch t is Π_{tk}, and we let $\Pi_t = \mathrm{diag}(\Pi_{t1}, \ldots, \Pi_{tK})$ collect the assignment probabilities in batch t. Observations assigned to arm k have mean outcome μ_k, collected in $\mu = [\mu_1, \ldots, \mu_K]'$, and variance σ_k^2, collected in $\Sigma = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_K^2)$. Once the experiment ends, either at an experimenter-selected stopping time ($T < T_0$) or upon reaching the last batch ($T = T_0$), we are interested in conducting inference on some linear combination of the arm means, $\eta'\mu$, where η may depend on the data.

The iterative nature of the experiment provides multiple opportunities for adaptive, data-driven decision-making. First, treatment assignments may be selected using bandit algorithms, which set Π_t using the results in batches $1, \ldots, t-1$. For instance, the experimenter might assign more units to arms which have produced good outcomes in the past, as in Thompson sampling. Second, the experimenter may choose whether to continue the experiment, and thus the observed number of batches T, based on how the results are unfolding. For instance, the experimenter might stop the experiment early if the results are either highly promising or overly discouraging. Third, the results of the experiment may also inform what objects the experimenter targets for inference, and hence the linear combination coefficients η. For instance, the experimenter might be interested in inference on average outcomes under an arm that performed especially well in the experiment.

Each of these forms of adaptivity presents challenges for inference. Zhang et al. (2020) and Hadad, Hirshberg, Zhan, Wager and Athey (2021) highlight that standard inference procedures that ignore the adaptive choice of Π_t can lead to invalid inference (e.g. under-coverage for confidence sets) in adaptive experiments, and propose alternative inference procedures which are valid in large samples provided η and T are fixed in advance. Similarly, the large and growing literature on anytime-valid inference is motivated by the long-standing observation that adaptively chosen stopping times T can lead to arbitrarily poor performance for conventional inference procedures, and proposes alternatives which guarantee performance in settings where η is fixed (see Ramdas, Grünwald, Vovk and Shafer (2023) for a recent review). Finally the large literatures on post-selection and selective inference (e.g. Berk, Brown, Buja, Zhang and Zhao, 2013; Fithian, Sun and Taylor, 2017; Andrews, Kitagawa and McCloskey, Forthcoming) highlight that the data-driven choice of a target parameter can invalidate standard inference procedures, and proposes valid alternatives focused primarily on static settings or, equivalently, settings where Π_t and T are fixed in advance.

In the batched setting we consider, there is a simple procedure available which ensures valid inference for many different choices of $(T, \Pi_{1:T}, \eta) = (T, \Pi_1, \ldots, \Pi_T, \eta)$, which we will collectively shorthand as "the experimental design." So long as the experimental design depends only on data observed up to period $T-1$—requiring, for instance, that decisions to stop the experiment be made one period in advance—inference based on the last batch of the experiment is guaranteed to be valid. This "last-batch-only" approach is in one sense quite restrictive, since it discards information from all but the last batch of the experiment. At the same time, last-batch only inference is quite flexible in that it requires no knowledge of how $(T, \Pi_{1:T}, \eta)$ are chosen, other than that (T, η) must be chosen one batch in advance. Consequently, this approach can accommodate a very wide variety of experimental designs, including ones where decisions are made adaptively by human decision-makers whose preferences and decision rules are not fully understood.

This paper examines the extent to which it is possible to improve last-batch-only inference while continuing to allow a very flexible class of experimental designs. In settings where we desire valid inference (e.g. coverage $1-\alpha$) conditional on $(T, \Pi_{1:T}, \eta)$ and there are no restrictions on how these are chosen, we prove that it is impossible to improve upon last-batch-only inference. We also observe, however, that many approaches for selecting $(T, \Pi_{1:T}, \eta)$ depend only on the *contrasts* between different arms, and are unaffected if we increase the average outcome for all arms by a fixed amount. This is the case, for instance, whenever the experimenter's decisions depend only on the difference in outcomes relative to a fixed treatment arm. We term experimental designs with this property "location-invariant," and show that in the class of location-invariant experiments, a sufficient statistic for μ is given by the outcomes in the last batch together with a scalar summarizing information from earlier batches. Estimation of $\eta'\mu$ based on these statistics is a simple generalized least squares problem, and dominates inference based on the last batch alone.

The class of location-invariant experimental designs is large, and in many contexts we know much more about $(T, \Pi_{1:T}, \eta)$ than simply that they are location-invariant. When we have such restrictions, we can exploit them to design more precise inference procedures. To illustrate, we consider settings where the rules used to determine $(T, \Pi_{1:T}, \eta)$ are fully known and can be expressed in terms of a finite number of linear-in-data inequalities (that is, settings where these $(T, \Pi_{1:T}, \eta)$ depend on the experimental results only through a collection of polyhedral events). This holds, for instance, when $\Pi_{1:T}$ is constructed using an ε-greedy algorithm and η selects the best-performing arm. In such settings, we characterize optimal median-unbiased estimators and equal-tailed confidence intervals conditional on $(T, \Pi_{1:T}, \eta)$.

Our optimality results are derived under the assumption that the batch-arm means are exactly normally distributed in finite samples. Hirano and Porter (2023) show that this corresponds to the limit experiment for batched bandit experiments. To ensure that our results for location-invariant experiments deliver valid inference more generally, in the appendix we prove that our procedures for location-invariant experimental designs are uniformly asymptotically valid across a large class of data generating processes with non-normal data.

Section 2 formally introduces the problem we consider in the context of the normal model, and proves the optimality of last-batch-only inference absent further restrictions on $(T, \Pi_{1:T}, \eta)$. Section 3 introduces the class of location-invariant experiments and derives the sufficient statistic for μ in this class. Section 4 considers optimal conditional inference when $(T, \Pi_{1:T}, \eta)$ depends on the data only through a collection of polyhedral events. Finally Section 5 provides simulation evidence on the performance of our procedures.

Proofs and additional results are provided in the appendix. In particular, Appendix A includes asymptotic results justifying the finite-sample normal approximation in the main text. Appendix B proves results stated in the main text.

2. Problem setup

Let n_t denote the total number of observations in batch t, $n = \sum_{t=1}^{T} n_t$ the total number of observations, and $c_t = n_t/n$ the fraction of observations assigned in batch t. We write X_{tk} for the average outcome among those units assigned to arm k in batch t and X_t for the vector of these means, $X_t = [X_{t1}, \ldots, X_{tK}]'$. We assume that the assignment probabilities Π_t depend on the data only through $X_{1:t-1} = (X_1, ..., X_{t-1})$.

Since X_t is a vector of sample means, when batch sizes are large the central limit theorem implies that X_t is approximately Gaussian conditional on $\Pi_t, X_{1:t-1}$, with $X_{tk} \mid \Pi_t, X_{1:t-1} \sim \mathcal{N}(\mu_k, \frac{\sigma_k^2}{n_t \Pi_{tk}})$. For V_t the diagonal matrix with k^{th} diagonal element $\frac{\sigma_k^2}{c_t \Pi_{tk}}$ we thus have that, approximately,

$$X_t \mid \Pi_t, X_{1:t-1} \sim \mathcal{N}\left(\mu, \frac{1}{n} V_t\right), \tag{1}$$

The results in the main text take (1) as exact and treat Σ (and thus V_t) as known, since the variances Σ are consistently estimable. Hirano and Porter (2023) show that this corresponds to the limit experiment for the batched bandit experiment under mild conditions. Moreover, in Appendix A we verify that the finite-sample analogue of our procedure converges to its counterpart under (1), uniformly over a large class of data-generating processes. For analytical convenience, we maintain in the main text that $\Pi_{tk} > 0$ almost surely for all $t \in [T], k \in [K]$. This is again relaxed by our asymptotic results in Appendix A.

2.1. **Inference problem.** Suppose that we observe data from batches $1, \ldots, T$ of the Gaussian experiment (1), where the stopping decision is based on information available at $T-1$ (i.e. $\mathbb{1}(T \leq t)$ is measurable with respect to $X_{1:t-1}$ for all t). We are interested in inference on $\eta'\mu$, where the target parameter is again determined by information available at $T-1$, $\eta = \eta(X_{1:T-1}, T)$.

The restriction that stopping decisions and target parameters be determined by $X_{1:T-1}$ implies that inference based only on the last batch is valid, in the sense that the z-statistic confidence interval $\eta' X_T \pm z_{1-\frac{\alpha}{2}} \sigma_\eta$ (where $z_{1-\frac{\alpha}{2}}$ is the $1-\frac{\alpha}{2}$ quantile of a standard normal distribution and $\sigma_\eta^2 = \frac{1}{n} \eta' V_T \eta$) has correct conditional coverage

$$\mathrm{P}\left(\eta'\mu \in \left[\eta' X_T \pm z_{1-\frac{\alpha}{2}} \sigma_\eta\right] \mid T, \Pi_{1:T}, \eta\right) = 1 - \alpha.$$

Moreover, $\eta' X_T$ is unbiased for $\eta'\mu$ even conditional on T, $\Pi_{1:T}$, and η

$$\mathbb{E}[\eta' X_T \mid T, \Pi_{1:T}, \eta] = \eta'\mu.$$

The first question we consider is whether we can construct a confidence set C_η that improves on last-batch-only inference while maintaining the same flexibility. If η can be an arbitrary function $\eta(\cdot)$ of the history $X_{1:T-1}$ and we wish to maintain inferential validity conditional on the stopping time, assignment probability, and target parameter

$$\mathrm{P}\left(\eta'\mu \in C_\eta(\alpha) \mid T, \Pi_{1:T}, \eta\right) \geq 1 - \alpha \text{ almost surely}, \quad \alpha \in (0, 1). \tag{2}$$

then the answer turns out to be no.

Lemma 2.1. *Suppose (2) holds for all measurable $\eta(\cdot)$, then for all fixed $\eta \in \mathbb{R}^K$*

$$\mathrm{P}\left(\eta'\mu \in C_\eta(\alpha) \mid X_{1:T-1}\right) \geq 1 - \alpha \textit{ almost surely.}$$

Lemma 2.1 shows that if $\eta(\cdot)$ is entirely unrestricted, then any confidence set which is valid conditional on $(T, \Pi_{1:T}, \eta)$ must also be valid conditional on $X_{1:T-1}$. In fact, simply requiring conditional coverage given η is already sufficient to obtain this conclusion. This effectively means we cannot use information from the first $T-1$ batches of data, so it is impossible to improve on the power of last-batch-only inference without losing some degree of robustness.

To make progress, we introduce two relaxations of (2). First, we observe that many bandit algorithms are *location-invariant*. That is, they have the property that $\Pi_t(X_{1:t-1})$ is invariant to adding a constant h to every batch-arm mean X_{sk}. If the adaptive choices of inferential target and stopping time are similarly location-invariant, as is often the case, then we can condition on less information, and construct procedures that dominate using solely the last batch. Such a procedure does not require knowledge of the precise allocation algorithm nor of $\eta(\cdot)$ and T, beyond location-invariance.

Second, if the experimental design depends solely on a lower-dimensional but known set of statistics, then we can also design optimal conditional inference procedures conditioning on these statistics. We show that such procedures are particularly tractable for a large class of discrete assignment algorithms that we call polyhedral algorithms. Since these procedures require less stringent conditioning, they are more powerful than conditional inference procedures that only use location-invariance.

3. **Conditional inference for location-invariant algorithms**

This section focuses on location-invariant assignment algorithms, choices of target parameter, and stopping times. We show that there is a simple conditional procedure that assumes only location-invariance and improves upon last-batch-only inference. To state these results, we first formally define what we mean by location-invariance.

Definition 3.1. A function $f(X_1, \ldots, X_t)$ is *location-invariant* if

$$f(X_1 + h1_K, \ldots, X_t + h1_K) = f(X_1, \ldots, X_t)$$

for all $(X_1, \ldots, X_t) \in (\mathbb{R}^K)^t$ and $h \in \mathbb{R}$, where $1_K \in \mathbb{R}^K$ is the vector of ones.

Definition 3.2. An assignment algorithm is *location-invariant* if each batchwise probability $\Pi_t = \Pi_t(X_1, \ldots, X_{t-1})$ is location-invariant.

Definition 3.3. A stopping time T is *location-invariant* if whether batch t is the last batch is determined by $X_{1:t-1}$ via a location-invariant function. That is, $\mathbb{1}(T > t) = \Xi_t(X_1, \ldots, X_{t-1})$ and $\Xi_t(\cdot)$ is location-invariant for all $t = 1, \ldots, T_0$.

Definition 3.4. η is location-invariant if $\eta = \eta(X_1, \ldots, X_T; T)$ and, for all $t = 1, \ldots, T_0$, each $\eta(X_1, \ldots, X_t; t)$ is location-invariant in its first t arguments.

For concreteness, let us first introduce two leading location-invariant assignment algorithms. To do so, let

$$W_t = \left(\sum_{s=1}^{t} V_s^{-1}\right)^{-1} \sum_{s=1}^{t} V_s^{-1} X_s$$

be the inverse-variance weighted batch-arm means, and let

$$\Omega_t = \frac{1}{n}\left(\sum_{s=1}^{t} V_s^{-1}\right)^{-1} = \frac{\Sigma}{n}\left(\sum_{s=1}^{t} c_s \Pi_s\right)^{-1}$$

be the posterior variance of $\mu \mid W_t$ under a flat prior.

Example 3.5 (Thompson sampling)**.** Let $Q(\nu, \Lambda)$ denote the vector of Gaussian orthant probabilities

$$Q_k(\nu, \Lambda) = \mathrm{P}_{X \sim \mathcal{N}(\nu, \Lambda)}\left(X_k \geq \max_{\ell} X_\ell\right). \tag{3}$$

Consider a Bayesian experimenter with a flat prior for μ. The Bayesian's posterior distribution for μ after observing $X_{1:t}$ is

$$\mu \mid X_{1:t} \sim \mathcal{N}\left(W_t, \Omega_t\right).$$

Thompson sampling then sets $\Pi_{t+1} = Q\left(W_t, \Omega_t\right)$. Since $Q(\nu + h, \Lambda) = Q(\nu, \Lambda)$, this algorithm is location-invariant. ■

Example 3.6 (ε-greedy)**.** For some $0 < \varepsilon < 1/K$, the algorithm sets $\Pi_{tk} = 1 - (K-1)\varepsilon$ if W_{tk} is the largest entry in W_t. Otherwise, the algorithm sets $\Pi_{tk} = \varepsilon$. This algorithm is location-invariant since it only involves ordinal comparisons of W_{tk}. ■

Location-invariance implies that for each $t \in \{1, ..., T_0\}$, $(\mathbb{1}(T \le t), \Pi_{1:t}, \eta)$ is measurable with respect to batch-arm differences against X_{1K}:

$$\Delta X_{1:t-1} \equiv \{X_{tk} - X_{1K} : s = 1, \ldots, t-1; k = 1, \ldots, K-1\}.$$

Consider the following statistic which represents some information from the first $T-1$ batches "left over" from conditioning on $(T, \Delta X_{1:T-1})$:

$$L = \sum_{t=1}^{T-1} \sum_{k=1}^{K} c_t \frac{\Pi_{tk}}{\sigma_k^2} X_{tk}$$

It turns out that L is the *only* information left over from conditioning, in the sense that (L, X_T) is sufficient for μ with respect to the conditional distribution of the data.

Theorem 3.7. *Under the preceding setup, assuming $\Pi_t > 0$ for all t almost surely,*

$$\begin{bmatrix} L \\ X_T \end{bmatrix} \mid (T, \Delta X_{1:T-1}) \sim \begin{bmatrix} L \\ X_T \end{bmatrix} \mid (\Pi_{2:T}, T) \sim \mathcal{N}\left(\begin{bmatrix} \lambda'\mu \\ \mu \end{bmatrix}, \begin{bmatrix} \frac{\lambda' 1}{n} & 0 \\ 0 & \frac{1}{n} V_T \end{bmatrix} \right), \tag{4}$$

where $\lambda = (\lambda_1, \ldots, \lambda_K)'$ and $\lambda_k = \frac{1}{\sigma_k^2} \sum_{t=1}^{T-1} \frac{n_t}{n} \Pi_{tk}$. Moreover, (L, X_T) is sufficient for μ with respect to the conditional distribution $X_{1:T} \mid (T, \Delta X_{1:T-1})$.

Therefore, we can base inference for μ on (4). Optimal inference in (4) is also optimal conditional inference given the finer information set $\Delta X_{1:T}$. The minimal sufficient statistic for μ under (4) is the weighted least-squares coefficient

$$S^\star = X_T + \frac{V_T \lambda}{\lambda' 1 + \lambda' V_T \lambda} (L - \lambda' X_T) \tag{5}$$

where

$$S^\star \mid (\Delta X_{1:T-1}, T) \sim \mathcal{N}\left(\mu, \frac{1}{n} \left(V_T - \frac{V_T \lambda \lambda' V_T}{\lambda' 1 + \lambda' V_T \lambda} \right) \right). \tag{6}$$

Optimal conditional inference for $\eta'\mu$ is thus simply based on the Gaussian statistic $\eta' S^\star$.

Appendix A formally connects these results in the Gaussian model to the finite sample setting. We show that the finite sample analogue of $\eta' S^\star$ has asymptotic behavior analogous to (6) and produces confidence intervals that cover the random parameter $\tau = \eta'\mu$ conditional on $(T, \Pi_{1:T}, \eta)$. Our asymptotic results are valid uniformly over a class of data-generating processes and also allow for certain arbitrarily small assignment probabilities, which may be of independent theoretical and practical interest.

3.1. **Gain relative to last-batch inference.** Theorem 3.7 shows that under location-invariance there is one usable piece of information beyond to the last batch, namely the statistic L. It is therefore interesting to understand to what extent this additional information helps us conduct inference on μ.

Let $r^{(c)} = V_T\lambda$. Note that $r^{(c)}$ collects the ratio of cumulative sample size and last period sample size for each arm: $r_k^{(c)} = \frac{\sum_{t=1}^{T-1} n_t \Pi_{tk}}{n_T \Pi_{Tk}}$. Similarly, let $r^{(t)} = r^{(c)} + 1$ collect the ratio of total sample size to last period sample size for each arm. Let $q = \sum_{t=1}^{T-1} n_t \Pi_t 1_K$ collect the cumulative sample size. Then we have that

$$\mathrm{Var}(S_k^\star \mid \Delta X_{1:T-1}, T) = \underbrace{\frac{V_T}{n}}_{\mathrm{Var}(X_T \mid \Pi_T)} \left(I - \frac{q\Sigma^{-1}(r^{(c)})'}{q'\Sigma^{-1} r^{(t)}} \right).$$

Thus, the reduction in variance relative to using solely the last batch depends on the alignment of η to the matrix $\frac{q\Sigma^{-1}(r^{(c)})'}{q'\Sigma^{-1} r^{(t)}}$, which collects information about variance-weighted relative sample sizes. The relative improvement is greatest when η is proportional to λ, since then L provides the greatest amount of information.

3.2. **Conditioning on $\Delta X_{1:T-1}$ in Thompson sampling.** We have so far studied inference conditional on all of the differences $\Delta X_{1:T-1}$. While this allows for any location-invariant form of selection, one might reasonably wonder if this conditioning is excessive. If we had more restrictions on how $(T, \Pi_{1:T}, \eta)$ were generated and our only goal were to ensure conditional coverage (2), perhaps it would suffice to condition on a coarser set of statistics, which could let us obtain higher power. The next section explores this possibility.

Before doing so, however, we show that conditioning on on $\Delta X_{1:T-1}$ is sometimes necessary for conditional coverage (2), even when the rules generating $(T, \Pi_{1:T}, \eta)$ are perfectly known. Specifically, we show in the case of Thompson sampling (Example 3.5), for fixed T and η, efficient conditional inference given $\Delta X_{1:T-1}$ is also efficient given $\Pi_{1:T}$. This is because $\Pi_{1:T}$ in Thompson sampling contains sufficiently rich information about $\Delta X_{1:T-1}$.

Proposition 3.1. *Under the setup of Theorem 3.7, suppose, for all t, $\Xi_t(X_1, \ldots, X_{t-1})$ is measurable with respect to $\Pi_{2:t}$. Under Thompson sampling, (L, X_T) is sufficient for μ with respect to the conditional distribution $X_{1:T} \mid \Pi_{1:T}, T$.*

Proposition 3.1 shows that (L, X_T) is sufficient for μ with respect to the distribution $X_{1:T} \mid (\Pi_{2:T}, T)$. Thus, optimal inference based on (L, X_T) is optimal for $\eta'\mu$. The proof relies on inversion results from the econometrics of discrete choice (Hotz and Miller, 1993; Norets and Takahashi, 2013) in order to show that $\Pi_{1:T}$ reveals sufficiently rich information on $\Delta X_{1:t-1}$.

4. Conditional inference in the Gaussian model

If we have additional knowledge of the allocation algorithm as well as the inferential target and stopping time, then we can design optimal conditional inference procedures that condition on less information—and are hence more powerful.

Let us first introduce a general recipe for constructing optimal conditional procedures by analyzing the likelihood of the observed data. Suppose we wish to condition on growing information sets $\mathcal{F}_{T-1}(X_{1:T-1})$ such that the experimental design $(\mathbb{1}(T \leq t), \Pi_{1:t}, \eta(X_1, \ldots, X_t; t))$ is measurable with respect to $\mathcal{F}_{t-1}$. The law of the observed data, for a given stopping time T', is an exponential family:

$$\begin{aligned} p_\mu(X_{1:T} \mid T = T', \mathcal{F}_{T'-1}) &= \frac{1}{p(\mathcal{F}_{T'-1})} \prod_{t=1}^{T'} \frac{1}{\sqrt{(2\pi)^k \det(V_t)}} \exp\left(-\frac{n}{2}(X_t - \mu)' V_t^{-1} (X_t - \mu)\right) \\ &= \exp\left(n\mu' \sum_{t=1}^{T'} V_t^{-1} X_t\right) h(X_{1:T'}) g(\mathcal{F}_{T'-1}, \mu) \end{aligned} \tag{7}$$

where $h(\cdot)$ does not depend on μ. Thus, a sufficient statistic for μ with respect to the law of $(X_{1:T} \mid T, \mathcal{F}_{T-1})$ is

$$S = n \sum_{t=1}^{T} V_t^{-1} X_t.$$

For any $\tau = \eta'\mu$, where without loss of generality $\|\eta\| = 1$, given an orthogonal matrix Q whose first row is η we can further partition

$$S'\mu = S'Q'Q\mu = (\eta'S)(\eta'\mu) + (\eta_\perp S)'(\eta_\perp \mu) \equiv U\tau + U_\perp' \tau_\perp.$$

By the results in Pfanzagl (1994), optimal inferences for τ, conditional on $\mathcal{F}_{T-1}$,[1] are based on the distribution of $U \mid U_\perp, \mathcal{F}_{T-1}$, which depends solely on τ and is stochastically increasing in τ. The optimal α-quantile unbiased estimator for τ, based on observations $(u, u_\perp, \mathcal{F}_{T-1}(X_{1:T-1}))$, can be computed by $\hat{\tau}_\alpha$ such that

$$\mathrm{P}_{\hat{\tau}_\alpha}(U \leq u \mid U_\perp = u_\perp, \mathcal{F}_{T-1}(X_{1:T-1})) = \alpha.$$

The optimal $(1-\alpha)$ equal-tailed confidence interval for τ is $[\hat{\tau}_{\alpha/2}, \hat{\tau}_{1-\alpha/2}]$.

If choices of inferential target and stopping time can be expressed in terms of linear inequalities in $X_{1:T-1}$, then the above recipe is particularly tractable. Like our results in Section 3, the optimal conditional procedures dominate using only the last batch.

Remark 4.1 (Sufficient statistics)**.** Notably, this procedure is based on $(U, U_\perp)$, which are sufficient for μ with respect to the law of $X_{1:T} \mid \mathcal{F}_{T-1}, T$. In contrast, we notice that the procedure in Zhang et al. (2020) is not based on the sufficient statistic. Appendix D shows that it can be written as a linear function of the sufficient statistics and an independent

[1] Adusumilli (2023) studies Neyman-Pearson-style testing with the likelihood (7).

Gaussian noise term. The natural estimator for τ from Zhang et al. (2020) (the mid point of the confidence interval) is thus dominated by its conditional expectation given the sufficient statistics (i.e. Rao–Blackwellization). Moreover, we observe that the confidence intervals in Zhang et al. (2020) is also not bet-proof (Müller and Norets, 2016), in the sense that there exists an event of the data with positive probability on which the confidence interval undercovers uniformly over $\mu \in \mathbb{R}^K$. ■

4.1. **Conditional inference for polyhedral algorithms.** We consider a class of algorithms such that the support of $\mathcal{F}_{T-1} \equiv (\Pi_{2:T}, T, \eta)$ is finite. For every (π, t_0, h) in its support, there exist a conformable matrix $A(\mathcal{F}_{T-1})$ and vector $b(\mathcal{F}_{T-1})$ such that for all μ,

$$
\begin{aligned}
&\mathrm{P}_\mu \left[(\Pi_{2:T}, T, \eta) = (\pi, t_0, h) \cap A(\mathcal{F}_{t_0-1})\mathbf{X}_{1:t} \leq b(\mathcal{F}_{t_0-1})\right] \\
&= \mathrm{P}_\mu \left[(\Pi_{2:T}, T, \eta) = (\pi, t_0, h)\right] \\
&= \mathrm{P}_\mu \left[A(\mathcal{F}_{t_0-1})\mathbf{X}_{1:t_0} \leq b(\mathcal{F}_{t_0-1})\right].
\end{aligned}
$$

As an example, suppose T, η are fixed, but at each batch t, the assignment algorithm ϵ-greedy (Example 3.6) favors the arm k with the highest cumulative arm mean $W_{(t-1),k}$. Each realization of $\pi_{2:T}$ then corresponds to a sequence of ϵ-greedy winners $(k_1, \ldots, k_{T-1})$. Thus, up to a measure-zero event of ties, the sequence of winners $(k_1, \ldots, k_{T-1})$ is equivalent to the inequalities

$$
\{W_{t,k_t} \geq W_{t,\ell} : t = 1, \ldots, T-1; \ell = 1, \ldots, K\}.
$$

Since $W_{t,\ell}$ is linear in $X_1, \ldots, X_t$, we may represent these inequalities as the polyhedron $A(\mathcal{F}_{T-1})\mathbf{X}_{1:T} \leq b(\mathcal{F}_{T-1})$.

Following our conditional inference strategy, optimal inference for $\tau = \eta'\mu$ depends on the distribution

$$
(U \mid U_\perp, \mathcal{F}_{T-1}) \sim (U \mid U_\perp, A(\mathcal{F}_{T-1})\mathbf{X}_{1:T} \leq b(\mathcal{F}_{T-1}))
$$

where $U = \eta' S$, $U_\perp = \eta_\perp S$, and $S_k = \sum_{t=1}^{T} \frac{n_t \Pi_{tk}}{\sigma_k^2} X_{tk}$. The following theorem makes explicit this distribution, and, in particular, its dependence on μ only through τ.

Theorem 4.2. *Assume, without loss of generality, that $\|\eta\| = 1$. Define $R = n \sum_{t=1}^{T} V_t^{-1}$. Let $G = 1_T' \otimes I_K = [I_K, \ldots, I_K]$ be a $K \times TK$ matrix. Let $G_\perp$ be a $(T-1)K \times TK$ matrix whose rows are unit vectors that are orthogonal to each other and to the rows of G. Let $\mathbf{V} = \frac{1}{n}\operatorname{diag}(V_1, \ldots, V_T)$ be a $TK \times TK$ diagonal matrix. Let $\tilde{G}, \tilde{G}_\perp$ partition the inverse of the following matrix:*

$$
\begin{bmatrix} G\mathbf{V}^{-1} \\ G_\perp \end{bmatrix}^{-1} = \left[\tilde{G}, \tilde{G}_\perp\right].
$$

where $\tilde{G}$ is $TK \times K$. Let $c = \frac{1}{\eta' R^{-1} \eta}$. Then

$$\left(\frac{U}{c} \mid U_\perp = u_\perp, A(\mathcal{F}_{T-1})\mathbf{X}_{1:T} \leq b(\mathcal{F}_{T-1})\right) \sim (Z_1 \mid M(\mathcal{F}_{T-1})Z \leq m(\mathcal{F}_{T-1}))$$

where the right-hand side follows the law

$$Z = (Z_1, Z_2')' \sim \mathcal{N}\left(\begin{bmatrix} \tau + k'u_\perp \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{c} & 0 \\ 0 & G_\perp \mathbf{V} G_\perp' \end{bmatrix}\right) \qquad k' = \frac{1}{c}\eta' G \mathbf{V}^{-1} G' \eta_\perp' (\eta_\perp G \mathbf{V}^{-1} G' \eta_\perp')^{-1}$$

The constraints are defined by

$$M(\mathcal{F}_{T-1}) = \left[cA(\mathcal{F}_{T-1})\tilde{G}\eta, A(\mathcal{F}_{T-1})\tilde{G}_\perp\right] \qquad m(\mathcal{F}_{T-1}) = b(\mathcal{F}_{T-1}) - A(\mathcal{F}_{T-1})\tilde{G}(I - \eta\eta')S.$$

At any realization of the data $X_{1:T}$, the constraints and parameters $(m(\mathcal{F}_{T-1}), M(\mathcal{F}_{T-1}), c, G_\perp \mathbf{V} G_\perp')$ are functions of known quantities. The law $(Z_1 \mid M(\mathcal{F}_{T-1})Z \leq m(\mathcal{F}_{T-1}))$ depends only on the parameter of interest τ.

For testing the point null hypothesis $H_0 : \tau = \tau_0$, we can draw from the corresponding distribution $(Z_1 \mid M(\mathcal{F}_{T-1})Z \leq m(\mathcal{F}_{T-1}))$ induced by the null value τ_0. The observed statistic is $z_1 = U/c$. Valid tests are then constructed by comparing z_1 to its distribution under the null. A level-$(1-\alpha)$ equal-tailed test, for instance, rejects H_0 when z_1 falls in the lower $\alpha/2$ or the upper $\alpha/2$ quantiles of the distribution $(Z_1 \mid M(\mathcal{F}_{T-1})Z \leq m(\mathcal{F}_{T-1}))$ under τ_0.

In terms of computation, we can efficiently draw from $(Z \mid M(\mathcal{F}_{T-1})Z \leq m(\mathcal{F}_{T-1}))$ via Gibbs sampling (Taylor and Benjamini, 2016),[2] since the conditional distribution of each coordinate of Z is a truncated Gaussian. Confidence intervals can be constructed by inverting tests of $H_0 : \tau = \tau_0$ for a range of τ_0. Computation of confidence intervals does not require drawing the distribution of Z for every candidate value of τ_0. Note that, since the likelihood ratio between two candidate values (τ_1, τ_0) is known,

$$\frac{p_{\tau_1}(z_1 \mid M(\mathcal{F}_{T-1})Z \leq m(\mathcal{F}_{T-1}))}{p_{\tau_0}(z_1 \mid M(\mathcal{F}_{T-1})Z \leq m(\mathcal{F}_{T-1}))} = \exp\left(c(\tau_1 - \tau_0)z_1\right),$$

samples from p_{τ_0} can be reweighted to obtain estimators for quantities under p_{τ_1}. As a result, inference based on Theorem 4.2 is computationally efficient, at least for moderately-sized (T, K).

5. Simulation evidence

We consider two simulated experiments. Both experiments have $T = 4, K = 3, \mu = [0, 0, 0]'$, $n_1 = n_2 = n_3 = n_4 = 200$. To demonstrate asymptotic validity of our inference procedures, we take the individual outcomes to be i.i.d. Rademacher random variables

[2] We thank Brad Ross for this suggestion.

Fixed Target
(inference on μ_3)

	Rejection rate	Median length	Median length relative to last
Leftover	0.051	0.449	0.881
ZJM	0.049	0.255	0.488
Last-only	0.048	0.524	1.000

Adaptive Target
(inference on arm with highest sample mean in the first $T - 1$ batches)

	Rejection rate	Median length	Median length relative to last
Leftover	0.052	0.312	0.917
ZJM	0.071	0.199	0.570
Last-only	0.051	0.350	1.000

Notes. 100,000 replications. Nominal 5% test. "Median length relative to last" takes the median of the ratio of confidence interval lengths relative to last-only, as opposed to the ratio of median lengths. □

TABLE 1. Thompson sampling experiment

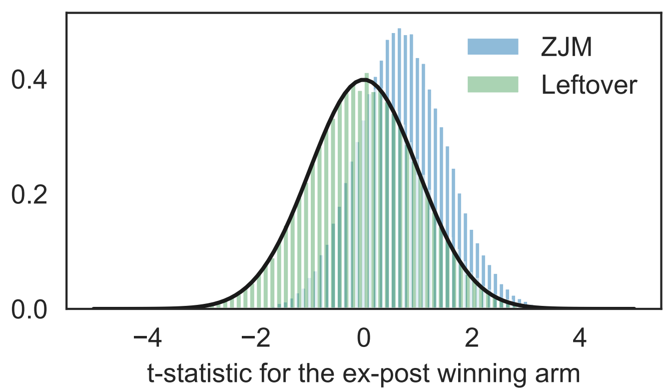


FIGURE 1. Distribution of t-statistics for adaptive target in the Thompson sampling experiment

and we plug in estimated versions of Σ in the following empirical exercises. We consider a Thompson sampling experiment and an ε-greedy experiment.[3]

For Thompson sampling, Table 1 displays coverage and length of different confidence intervals (Leftover is the procedure in Section 3 (Theorem A.2), ZJM is the procedure in

[3]The Thompson sampling algorithm *prunes* the last-batch assignment probabilities at 0.01 (See Example A.3), which is needed for our asymptotic results in Appendix A. The ε-greedy algorithm chooses $\varepsilon = 0.1$.

Fixed Target
(inference on μ_3)

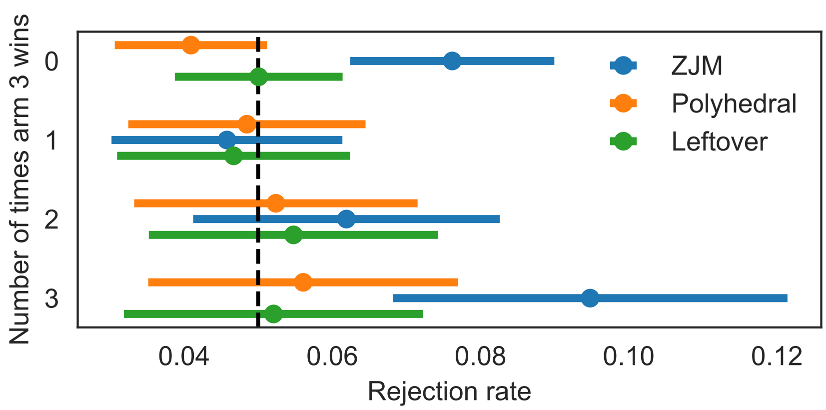
	Rejection rate	Median length	Median length relative to last
Polyhedral	0.049	0.368	0.622
Leftover	0.051	0.597	0.887
ZJM	0.052	0.271	0.428
Last-only	0.051	0.817	1.000

Adaptive Target
(inference on arm with highest sample mean in the first $T-1$ batches)

	Rejection rate	Median length	Median length relative to last
Polyhedral	0.048	0.244	0.789
Leftover	0.051	0.293	0.950
ZJM	0.067	0.203	0.654
Last-only	0.051	0.310	1.000

Notes. 100,000 replications. Nominal 5% test. "Median length relative to last" takes the median of the ratio of confidence interval lengths relative to last-only, as opposed to the ratio of median lengths. □

TABLE 2. ε-greedy experiment



Notes. Simultaneous 95% confidence intervals for coverage shown. □

FIGURE 2. Conditional coverage in ε-greedy experiment

Zhang et al. (2020),[4] and Last-only is the procedure that only uses the last batch). We compare the procedures over two setups, fixed target and adaptive target. In terms of

[4]That is, we use the estimated counterpart of the following pivotal quantity

$$Z_{\mathrm{ZJM}} = \sum_{t=1}^{T} V_t^{-1/2}(X_t - \mu) \sim \mathcal{N}(0, TI_K).$$

rejection rates, the conditional procedures control size for both setups, and ZJM does not control size for the adaptive setup (Figure 1 plots the distribution of t-statistics in the adaptive setup). In terms of length, the improvement in median length from using the leftover information is about 10% for this particular setup of the experiment, relative to only using the last batch, whereas the ZJM interval is 40–50% shorter than only using the last batch. This confirms that the improvement that the additional statistic L provides relative to last-batch-only is mild but non-trivial (a 10% improvement in length is comparable to a 20% increase in sample size). On average, we do sacrifice confidence interval length to maintain conditional validity, relative to the unconditionally valid procedure ZJM, and so the length difference between Leftover and ZJM can be viewed as the price of additional adaptivity.

Similarly, we additionally compare the behavior of the Polyhedral inference procedure (Theorem 4.2) for an ε-greedy experiment in Table 2. For inference on a fixed target, all procedures achieve their nominal size. The Polyhedral procedure generates 40% shorter confidence intervals (measured in terms of median length) than Last-only, which is more than double the improvement of Leftover. For inference on the adaptive target, we again see that the conditional procedures maintain nominal size, whereas ZJM does not.

Lastly, we plot conditional behavior of these procedures in Figure 2, where we condition on the number of times the inference target, arm 3, is the ε-greedy winner. We find that the conditional procedures indeed control conditional size, whereas ZJM over-rejects when arm 3 (for whose mean we perform inference) wins most and least often, and compensates for the over-rejection by under-rejecting on other sequences of ε-greedy winners.

6. **Conclusion**

This paper investigates inference conditional on the experimental design in batched adaptive experiments, and explores the potential of improving upon inference procedures that only use the last batch. For location-invariant experimental designs, we find there is scalar statistic, beyond the last-batch result, which is left over after conditioning. Using this additional statistic provides a free lunch improvement for statistical inference, relative to using the last batch alone. For polyhedral experimental designs, we characterize optimal conditional inference procedures and demonstrate their computational tractability.

References

Adusumilli, Karun, "Optimal tests following sequential experiments," *arXiv preprint arXiv:2305.00403*, 2023. 9

Andrews, Donald WK, Xu Cheng, and Patrik Guggenberger, "Generic results for establishing the asymptotic size of confidence sets and tests," 2011. 17, 30

Andrews, Isaiah, Toru Kitagawa, and Adam McCloskey, "Inference on winners," Technical Report Forthcoming. 2

Berk, Richard, Lawrence Brown, Andreas Buja, Kai Zhang, and Linda Zhao, "Valid post-selection inference," *Annals of Statistics*, 2013, *41* (2), 802–831. 2

Fithian, William, Dennis Sun, and Jonathan Taylor, "Optimal Inference After Model Selection," *arXiv*, 2017. 2

Hadad, Vitor, David A Hirshberg, Ruohan Zhan, Stefan Wager, and Susan Athey, "Confidence intervals for policy evaluation in adaptive experiments," *Proceedings of the National Academy of Sciences*, 2021, *118* (15), e2014602118. 2

Hirano, Keisuke and Jack R. Porter, "Asymptotic Representations for Sequential Decisions, Adaptive Experiments, and Batched Bandits," 2023. 2, 3, 4

Hotz, V Joseph and Robert A Miller, "Conditional choice probabilities and the estimation of dynamic models," *The Review of Economic Studies*, 1993, *60* (3), 497–529. 8, 26

Müller, Ulrich K and Andriy Norets, "Credibility of confidence sets in nonstandard econometric problems," *Econometrica*, 2016, *84* (6), 2183–2213. 10

Norets, Andriy and Satoru Takahashi, "On the surjectivity of the mapping between utilities and choice probabilities," *Quantitative Economics*, 2013, *4* (1), 149–155. 8, 26

Pfanzagl, Johann, *Parametric statistical theory*, Walter de Gruyter, 1994. 9

Rakočević, Vladimir, "On continuity of the Moore-Penrose and Drazin inverses.," *Matematichki Vesnik*, 1997, *49*, 163–172. 40

Ramdas, Aaditya, Peter Grünwald, Vladimir Vovk, and Glenn Shafer, "Game-theoretic statistics and safe anytime-valid inference," *Statistical Science*, 2023. 2

Taylor, J and Y Benjamini, "RestrictedMVN: multivariate normal restricted by affine constraints," *R package version*, 2016, *1.* 11

van der Vaart, Aad W and Jon Wellner, *Weak convergence and empirical processes: with applications to statistics*, Springer Science & Business Media, 1996. 34, 39

Vershynin, Roman, *High-dimensional probability: An introduction with applications in data science*, Vol. 47, Cambridge university press, 2018. 40

Zhang, Kelly, Lucas Janson, and Susan Murphy, "Inference for batched bandits," *Advances in Neural Information Processing Systems*, 2020, *33*, 9818–9829. 2, 9, 10, 13, 16, 45

Appendix for "Optimal Conditional Inference in Adaptive Experiments"

Jiafeng Chen and Isaiah Andrews

September 22, 2023

Contents

Appendix A. Uniform asymptotics

While the main text focuses on the finite-sample normal model, it is motivated by a large-batch asymptotic approximation. In this section we make this precise by explicitly showing that the finite-sample analogue of the inference procedure following Theorem 3.7 has desired asymptotic validity in a uniform sense.

To state these results, we introduce the finite-sample setting formally. Fix a sequence $(n, n_1, \ldots, n_{T_0})$ indexed by n such that $n_t/n \to c_t \in (0, 1)$ and $\sum_t n_t = n$. Let $\mathcal{I}_t \subset [n]$ collect the indicies of the n_t individuals in batch t. Prior to each batch, assignment probabilities $\Pi_{tn} = \operatorname{diag}(\Pi_{t1n}, \ldots, \Pi_{tKn})$ are determined by the assignment algorithm, as a function of past data. In batch t, the n_t participants are assigned to one of K arms. The indicator for individual i's assignment is $D_i = [D_{i1}, \ldots, D_{iK}]'$, which follows a categorical distribution with probabilities Π_{tn}. The D_i's are independent across i. Let $N_{tn} = [N_{t1n}, \ldots, N_{tKn}]' = \sum_{i \in \mathcal{I}_t} D_i$. Let $\hat{\Pi}_{tkn} = N_{tkn}/n_t$ be the realized frequency of samples to arm k in batch t. Let $\hat{\Pi}_{tn} = \operatorname{diag}(\hat{\Pi}_{t1n}, \ldots, \hat{\Pi}_{tKn})$.

Let X_i^{obs} be the outcome observed for i and let

$$X_{tkn} = \frac{1}{N_{tkn} \vee 1} \sum_{i \in \mathcal{I}_t} D_{ik} X_i^{\text{obs}}$$

be the batch-arm mean. The definition for X_{tkn} builds in the convention that an empty mean is zero. Let

$$\hat{\sigma}^2_{tkn} = \frac{\sum_{s=1}^t \sum_{i \in \mathcal{I}_s} D_{ik}(X_i^{\text{obs}} - W_{t,k,n})^2}{\sum_{s=1}^t N_{skn}} \tag{A.1}$$

be an estimate of the arm variance using data up until batch t. Let $\hat{\sigma}^2_{kn} = \hat{\sigma}^2_{Tkn}$. Let $\hat{\Sigma}_{tn} = \text{diag}(\hat{\sigma}^2_{t1n}, \ldots, \hat{\sigma}^2_{tKn})$ and let $\hat{\Sigma}_n = \text{diag}(\hat{\sigma}^2_{1n}, \ldots, \hat{\sigma}^2_{Kn})$.

Let the realized stopping time be T_n and the realized direction of inference be η_n. Suppose the experimenter wishes to conduct inference for a possibly data-dependent parameter $\tau_n = \eta_n' \mu$. We let

$$\lambda_n = \sqrt{n} \sum_{t=1}^{T_n - 1} \frac{n_t \hat{\Pi}_{tkn}}{n \hat{\sigma}^2_{kn}} \quad L_n = \sqrt{n} \sum_{t=1}^{T_n-1} \sum_{k=1}^K \frac{n_t \hat{\Pi}_{tkn}}{n \hat{\sigma}^2_{kn}} X_{tkn}$$

be scaled analogues of λ and L. Finally, let

$$S_n^\star = \left(\frac{\lambda_n \lambda_n'}{\sqrt{n} \lambda_n' 1} + \frac{n_{T_n}}{n} \hat{\Sigma}_n^{-1} \hat{\Pi}_{T_n,n} \right)^+ \left(\frac{\lambda_n}{\lambda_n' 1} \frac{L_n}{\sqrt{n}} + \frac{n_{T_n}}{n} \hat{\Sigma}_n^{-1} \hat{\Pi}_{T_n n} X_{T_n n} \right) \tag{A.2}$$

be an analogue of (5).[5] The estimator for τ_n is $\hat{\tau}_n = \eta_n' S_n^\star$, with estimated asymptotic standard error

$$\hat{\sigma}^2_{\tau,n} = \eta_n' \left(\frac{\lambda_n \lambda_n' / \sqrt{n}}{\lambda_n' 1} + \frac{n_{T_n}}{n} \hat{\Sigma}_n^{-1} \hat{\Pi}_{T_n,n} \right)^+ \eta_n.$$

Our main asymptotic result is that, uniformly over a class $\mathcal{P}$ in the sense of Andrews, Cheng and Guggenberger (2011), $\sqrt{n} \hat{\sigma}^{-1}_{\tau,n} (\eta_n' S_n^\star - \tau_n) \xrightarrow{d} Z$ where Z is conditionally standard Gaussian given the distributional limits of the experimental design $(T_n, \Pi_{1:T_n}, \eta_n)$. Thus, $(1-\alpha)$ two-sided confidence intervals $\text{CS}_n(\alpha) \equiv \eta_n' S_n^\star \pm \Phi^{-1}(1 - \alpha/2) \cdot \hat{\sigma}^2_{\tau,n}$ have exact asymptotic conditional size for the parameters τ_n:

$$\limsup_{n \to \infty} \sup_{P \in \mathcal{P}} \left| P\left(\tau_n \in \text{CS}_n(\alpha) \mid T_n = T, \eta_n = \eta\right) - \alpha \right| P(T_n = T, \eta_n = \eta) = 0. \tag{A.3}$$

Such a statement is predicated on assumptions about the class $\mathcal{P}$, the assignment algorithm for generating $\Pi_{1:T_0,n}$, T_n, and $\eta_n(\cdot)$. We assume that members of $\mathcal{P}$ have second moments bounded away from 0 and ∞ and uniformly bounded fourth moments.

[5] Here, to conform with our asymptotic results, we study the following parametrization of (4)

$$\begin{bmatrix} L \\ \sqrt{n_t} \Pi_T X_T \end{bmatrix} \mid \Pi_{2:T} \sim \mathcal{N}\left(\begin{bmatrix} \lambda' \mu \\ \sqrt{n_t} \Pi_t \mu \end{bmatrix}, \begin{bmatrix} \frac{\lambda' 1}{n} & 0 \\ 0 & \Pi_T \Sigma \end{bmatrix} \right)$$

which we again treat as a weighted least-squares problem for μ. (5) follows from applying the Sherman–Morrison identity in the case that $\Pi_T > 0$.

Assumption A.1. *$\mathcal{P}$ is a class of distributions satisfying the following conditions indexed by constants $C_1, C_2 > 0$: For all $P \in \mathcal{P}$,*

(1) For all k, $\frac{1}{C_1} \le \sigma_k^2(P) \le C_1$.[6]

(2) $\mathbb{E}_P[\|X\|^4] < C_2$ where $X \sim P$ is a vector of potential outcomes.

In this section, we consider $T_n, \eta_n, \Pi_{t+1,n}$ that depend on $X_{1:T_0}$ only through the cumulative mean vector

$$W_{tkn} = \sum_{s=1}^{t} \frac{n_s \hat{\Pi}_{skn}}{\sum_{r=1}^{t} n_r \hat{\Pi}_{rkn}} X_{skn}.$$

Our proof in Appendix C.6 considers weaker conditions.

In particular, let

$$W_{tn} = \left(\sum_{s=1}^{t} \Pi_{sn}\right)^{-1} \sum_{s=1}^{t} \Pi_{sn} X_{sn}$$

$$\Omega_{tn} = \hat{\Sigma}_n \left(\sum_{s=1}^{t} \frac{n_t}{n} \hat{\Pi}_{sn}\right)^{-1}$$

be the analogues of W_t, Ω_t defined in the main text. We assume that

$$\mathbb{1}(T_n > t+1) = \Xi_{t+1}\left(\sqrt{n} W_{tn}, \Omega_{tn}\right)$$

for some known function Ξ_t. Likewise, we assume that the assignment probability

$$\kappa_{t+1,n}\left(X_{1:t,n}, \hat{\Pi}_{1:t,n}, \hat{\Sigma}_{tn}, n_{1:t}\right) = \kappa_{t+1}\left(\sqrt{n} W_{tn}, \Omega_{tn}\right). \tag{A.4}$$

We assume that both Ξ_{tn} and $\kappa_{t+1,n}$ are location-invariant in W. Moreover, we require these functions to be suitably continuous, so that they converge weakly when their arguments converge weakly. All of the following assumptions are relative to some $\epsilon > 0$.

Definition A.1. Consider a function $g(w, \Omega)$, we say that g is *location-invariant and adequately continuous* (LIAC) if the following is true:

(1) For all $c \in \mathbb{R}$, $g(w + c1_K, \Omega) = g(w, \Omega)$

(2) For every nonempty $J \subset [K]$, every diagonal matrix Ω with entries within $[\epsilon, 1/\epsilon]$, Lebesgue-*almost* every $(w_j : j \in J) \in \mathbb{R}^{|J|}$, and every sequence $\mathbb{R}^K \ni w_m \to w_\infty$ and $\Omega_m \to \Omega$ where

$$w_{\infty,j} = \begin{cases} w_j, & \text{if } j \in J \\ -\infty & \text{otherwise.} \end{cases}$$

and Ω_m are positive-definite diagonal matrices, we have that

$$g(w_m, \Omega_m) \to g(w_\infty, \Omega). \tag{A.5}$$

[6]The upper bound is redundant given item (2), but we will keep it for convenience.

Assumption A.2. *For all $t \in \{2, \ldots, T_0 - 1\}$, $\Xi_t(\cdot)$ is LIAC, and $\Xi_t \leq \Xi_{t-1}$ almost surely. Moreover, $\Xi_1 = 1$ and $\Xi_{T_0} = 0$.*

Assumption A.3. *For all $t \in \{2, \ldots, T_0 - 1\}$,*

$$\kappa_{t+1} = \underbrace{\Xi_t \Xi_{t+1}}_{\mathbb{1}(T_n > t+1)} \kappa_{t+1,(0)} + \underbrace{(\Xi_t - \Xi_{t+1})}_{\mathbb{1}(T_n = t+1)} \kappa_{t+1,(1)} + \underbrace{(1 - \Xi_t)}_{\mathbb{1}(T_n \leq t)} \Pi_1. \tag{A.6}$$

and $\kappa_{t+1,(1)}, \kappa_{t+1,(0)}$ are LIAC functions of $\sqrt{n} W_{tn}, \Omega_{tn}$. Moreover, $\kappa_{t+1,(0),k}(w, \Omega) > 0$ when $w_k > 0$, and Π_1 is fixed and its entries are bounded below by ϵ.

The expression (A.6) means that κ_{t+1} takes the form of a contingency plan. If the experiment continues, then the next-batch probabilities are prescribed by $\kappa_{t+1,(0)}$. Otherwise, they are prescribed by $\kappa_{t+1,(1)}$. For analytical convenience, we say that $\kappa_{t+1} = \Pi_1$ when the experiment has already stopped, but this is not material since our statistic do not consider random variables beyond batch T_n.

For the pseudoinverse $\left(\frac{\lambda_n \lambda_n'}{\sqrt{n} \lambda_n' 1} + \frac{n_T}{n} \hat{\Sigma}_n^{-1} \hat{\Pi}_{Tn} \right)^+$ to converge weakly when its arguments do, we additionally require that the rank of $\frac{\lambda_n \lambda_n'}{\sqrt{n} \lambda_n' 1} + \frac{n_T}{n} \hat{\Sigma}_n^{-1} \hat{\Pi}_{Tn}$ is well-behaved. A sufficient condition is to assume that the range of the assignment probabilities of the last batch excludes $(0, \epsilon)$. This can be implemented by *pruning* away small probabilities and set them to zero, and redistributing the excess mass.

Assumption A.4. *$\kappa_{t+1,(1)}$'s range excludes $(0, \epsilon)$ on every coordinate.*

Lastly, we consider a simple form of η_n. We assume that η_n takes finitely many values, but they can be chosen with respect to some rich information.

Assumption A.5. *η_n takes finitely many values depending on T_n, the locations of the zero entries in $\Pi_{T_n,n}$, and the ordering of $W_{T_n,n}$: Let S_K be the set of permutations of K symbols,*

$$\eta_n = \sum_{\varsigma \in S_K : W_{T_n,n,\varsigma(1)} < \cdots < W_{T_n,n,\varsigma(K)}} \sum_{t=1}^{T_0} \sum_{E \subset [K]} \mathbb{1}\left(T_n = t, \{k : \Pi_{T_n,n,k} = 0\} = E\right) \eta_{\varsigma,t,E}.$$

Moreover, $\|\eta_{\varsigma,t,E,k}\| > \epsilon$ and $\eta_{\varsigma,t,E,k} = 0$ whenever $k \in E$.

We note that Assumptions A.2 to A.5 are stronger than necessary. We state weaker conditions as Assumptions C.1 to C.5, and verify that the weaker conditions are implications of the stronger ones. Under these assumptions, we formally justify (A.3).

Theorem A.2. *Under Assumptions A.1 to A.5, level-$(1-\alpha)$ two-sided confidence intervals $\mathrm{CS}_n(\alpha) \equiv \eta_n' S_n^\star \pm \Phi^{-1}(1-\alpha/2) \cdot \frac{\hat{\sigma}_{\tau,n}^2}{\sqrt{n}}$ have exact conditional asymptotic size for the parameters τ_n: For all η, T,*

$$\liminf_{n \to \infty} \inf_{P \in \mathcal{P}} \left| P\left(\tau_n \in \mathrm{CS}_n(\alpha) \mid \eta_n = \eta, T_n = T\right) - \alpha \right| P(\eta_n = \eta, T_n = T) = 0.$$

Another notable property of the asymptotics is that it does not require truncation of $\Pi_{2:T-1,n}$ and instead solely requires a pruning step for Π_{Tn}. This is because the statistic (A.2) can be written in terms of the scaled means $Y_{tkn} = \sqrt{c_t}\hat{\Pi}_{tkn}(X_{tkn} - \mu_k(P_n))$ instead of X_{tkn}. The statistics Y_{tkn} admit a central limit theorem uniformly over $\Pi_{tkn} \in [0,1]$.

A.1. Algorithms satisfying Assumptions A.3 and A.4.

Example A.3 (Pruned Thompson sampling)**.** Pruned Thompson sampling is the assignment algorithm of the form (A.6) where $\kappa_{t,(1)} = (\rho \circ Q)(\cdot,\cdot)$ and $\kappa_{t,(0)} = Q(\cdot,\cdot)$ for the Gaussian orthant probability Q defined in (3) and the pruning function

$$\rho_k(\pi) \propto \pi_k \mathbb{1}(\pi_k \geq \epsilon)$$

and $\sum_{k=1}^K \rho_k(\pi) = 1$.

Lemma A.1. *This procedure satisfies Assumptions A.3 and A.4.*

Proof. It suffices to check that (a) both $\rho \circ Q$ and Q are LIAC for the continuity statements, (b) $Q_k(v, \Omega) > 0$ when $v_k > 0$.

Note that (b) is immediately true. Now (a) for Q is immediately true as well, since Q is continuous at every input ν, Ω where Ω is a diagonal matrix with entries between ϵ and $1/\epsilon$ and $\nu \in [-\infty, 0]^K$.

Lastly, fix a nonempty $J \subset [K]$. For $\kappa_{t,(1)}$, for a fixed Ω, the discontinuities are of the form

$$\bigcup_{k \in J} \left\{\nu \in \mathbb{R}^{|J|} : Q_k(\nu^*, \Omega) = \epsilon, \nu_k^* = -\infty \text{ for } k \notin J, \nu_j^* = \nu_j\right\}.$$

Note that each entry of the union is measure zero with respect to $\mathbb{R}^{|J|}$. Hence, the (A.5) is satisfied for almost every ν. This verifies Assumption A.3. The pruning satisfies Assumption A.4 by construction. □

■

Example A.4 (ε-greedy)**.** Consider the algorithm satisfying (A.4) and

$$\kappa_{t,(1)}(\nu, \Omega) = \kappa_{t,(0)}(\nu, \Omega) = \begin{cases} 1 - \varepsilon, & \text{if } \nu_k \geq \max_\ell \nu_\ell \\ \varepsilon/(K-1) & \text{otherwise.} \end{cases}$$

for some $\varepsilon > 0$ (defined on ν where the largest entry is unique).

Lemma A.2. *This procedure satisfies Assumptions A.3 and A.4.*

Proof. For $\epsilon < \varepsilon/(K-1)$, this procedure satisfies Assumption A.4 by construction. For Assumption A.3, fix some J. Note that the discontinuities are of the form

$$\bigcup_{k \neq \ell, k \in J, \ell \in J} \left\{\nu : \nu_k = \nu_\ell \geq \max_{m \in [K]} \nu_m > -\infty\right\}.$$

Each entry of the union is measure zero with respect to $\mathbb{R}^{|J|}$. Hence, (A.5) is satisfied for almost every ν. Moreover, note that the assignment probabilities are never zero, and so the second part of Assumption A.3 is automatically satisfied. This verifies Assumption A.3. $\square$

$\blacksquare$

Appendix B. Proofs for results in the main text

B.1. Improvability.

Lemma 2.1. *Suppose (2) holds for all measurable $\eta(\cdot)$, then for all fixed $\eta \in \mathbb{R}^K$*

$$\mathrm{P}\left(\eta'\mu \in C_\eta(\alpha) \mid X_{1:T-1}\right) \geq 1-\alpha \text{ almost surely.}$$

Proof. Suppose not. Then there exists some event E and some η_0 where $\mathrm{P}(X_{1:T-1} \in E) > 0$ and

$$\mathrm{P}\left(\eta_0'\mu \in \hat{C}_{\eta_0}(\alpha) \mid X_{1:T-1} \in E\right) < 1-\alpha.$$

Let $\eta = \eta_0$ on E and $\eta = e_1$ on E^C. Then

$$\mathrm{P}\left(\eta'\mu \in \hat{C}_\eta(\alpha) \mid \eta(X_{1:T-1}) = \eta_0\right) < 1-\alpha$$

and $\mathrm{P}(\eta(X_{1:T-1}) = \eta_0) > 0$. This contradicts (2). $\square$

B.2. Leftover information L.

Theorem 3.7. *Under the preceding setup, assuming $\Pi_t > 0$ for all t almost surely,*

$$\begin{bmatrix} L \\ X_T \end{bmatrix} \mid (T, \Delta X_{1:T-1}) \sim \begin{bmatrix} L \\ X_T \end{bmatrix} \mid (\Pi_{2:T}, T) \sim \mathcal{N}\left(\begin{bmatrix} \lambda'\mu \\ \mu \end{bmatrix}, \begin{bmatrix} \frac{\lambda' 1}{n} & 0 \\ 0 & \frac{1}{n}V_T \end{bmatrix}\right), \tag{4}$$

where $\lambda = (\lambda_1, \ldots, \lambda_K)'$ and $\lambda_k = \frac{1}{\sigma_k^2}\sum_{t=1}^{T-1} \frac{n_t}{n}\Pi_{tk}$. Moreover, (L, X_T) is sufficient for μ with respect to the conditional distribution $X_{1:T} \mid (T, \Delta X_{1:T-1})$.

Proof. We first scale X_t by Π_t and use Lemma B.1—doing so mainly conforms with our subsequent asymptotic results. Recall that $\Pi_t = \operatorname{diag}(\Pi_{t1}, \ldots, \Pi_{tK})$, $\Sigma = \operatorname{diag}(\sigma_1^2, \ldots, \sigma_K^2)$, $V_T = \frac{1}{n}\frac{1}{c_T}\Pi_T^{-1}\Sigma$ where $c_t = n_t/n$. Note that we can represent conditioning on $\Delta X_{1:T-1}$ as conditioning on certain linear transformations $A\mathbf{X}_{1:T-1}$, where $\mathbf{X}_{1:T-1} = [X_1', \ldots, X_{T-1}']'$.

To apply Lemma B.1, define $Y_t = \sqrt{c_t}\Pi_t\sqrt{n}(X_t - \mu)$ and let $\mathbf{Y}_{1:t} = [Y_1', \ldots, Y_t']'$. When $\Pi_t > 0$, it suffices to study the joint distribution of L and Y_T conditional on $B\mathbf{Y}_{1:(T-1)}$, where

$$B = A\begin{bmatrix} \sqrt{nc_1}\Pi_1 & & \\ & \ddots & \\ & & \sqrt{nc_{T-1}}\Pi_{T-1} \end{bmatrix}^{-1}.$$

Note that since A takes pairwise differences, $A1_{(T-1)K} = 0$. Thus,

$$B \begin{bmatrix} \sqrt{c_t}\Pi_1 1_K \\ \vdots \\ \sqrt{c_{T-1}}\Pi_{T-1}1_K \end{bmatrix} = 0.$$

Note that we can define A_t, B_t to be the submatrices that only act on $Y_{1:t}$ such that $A = A_{T-1}$, $B = B_{T-1}$. Note that for all t, (B.1) is satisfied. Thus we can apply Lemma B.1 to show that

$$\begin{bmatrix} \sqrt{n}(L - \lambda'\mu) \\ Y_T \end{bmatrix} \mid T, B_{T-1}\mathbf{Y}_{1:T-1} \sim \mathcal{N}\left(0, \begin{bmatrix} \lambda' 1 & 0 \\ 0 & \Pi_T\Sigma \end{bmatrix}\right)$$

Therefore,

$$\begin{bmatrix} L \\ X_T \end{bmatrix} \mid T, A\mathbf{X}_{1:T-1} \sim \mathcal{N}\left(\begin{bmatrix} \lambda'\mu \\ \mu \end{bmatrix}, \begin{bmatrix} \lambda' 1/n & 0 \\ 0 & \frac{1}{n}V_T \end{bmatrix}\right)$$

Since the right-hand side of the above display depends only on $\Pi_{2:T}, T$,

$$\begin{bmatrix} L \\ X_T \end{bmatrix} \mid T, A\mathbf{X}_{:T-1} \sim \begin{bmatrix} L \\ X_T \end{bmatrix} \mid \Pi_{2:T}, T.$$

This proves (4).

To prove the "moreover" part, we show that the information in $\mathbf{X}_{1:T-1}$ can be recovered from L and $\Delta X_{1:T-1}$. Since $\Delta X_{1:T-1}$ and L are both linear transformations of $\mathbf{X}_{1:T-1}$, it suffices to show that the coefficients that generate $\Delta X_{1:T-1}$ and L are linearly independent. Precisely speaking, note that by construction, $L = \ell'\mathbf{X}_{1:T-1}$ where $AD\ell = 0$ (almost surely for some positive definite diagonal matrix D that depends on T). It suffices to show that the rows of A and ℓ span the whole space $\mathbb{R}^{K(T-1)}$. Since A is rank $K(T-1)-1$, it suffices to show that ℓ is not spanned by the rows of A.

The condition $AD\ell = 0$ is sufficient for this. This shows that ℓ is linearly independent with rows of A, since $D\ell \in \mathrm{Nul}(A) = \mathrm{Col}(A')^\perp$. Since $\ell' D\ell > 0$, $\ell \notin \left(\mathrm{Col}(A')^\perp\right)^\perp = \mathrm{Col}(A')$. Hence ℓ is not in the row space of A. Since A is rank-deficient by a single rank, we conclude that we can invert $L, A\mathbf{X}_{1:T-1}$ into $\mathbf{X}_{1:T-1}$.

Hence $(L, X_T, A\mathbf{X}_{1:(1-T)})$ maps into $X_{1:T}$, which is trivially sufficient for μ with respect to $X_{1:T} \mid T$. Since $(L, X_T, A\mathbf{X}_{1:(1-T)})$ is sufficient for μ with respect to $X_{1:T} \mid T$, (L, X_T) is sufficient with respect to $X_{1:T} \mid T, \Delta X_{1:T-1} \sim X_{1:T} \mid \Delta X_{1:T}$. This proves the "moreover" part. $\square$

Lemma B.1. *In the Gaussian model (where we allow $\Pi_{tk} = 0$ with positive probability), recall that $\Pi_t = \mathrm{diag}(\Pi_{t1}, \ldots, \Pi_{tK})$ and $\Sigma = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_K^2)$. Let $Y_t = \sqrt{c_t}\Pi_t\sqrt{n}\,(X_t - \mu)$ such that $Y_t \mid \Pi_t \sim \mathcal{N}(0, \Pi_t\Sigma)$. Let $\mathbf{Y}_{1:t} = [Y_1', \ldots, Y_t']'$. Let $B_t = B_t(\Pi_{1:t})$ be a matrix such* *that*

$$B_t \begin{bmatrix} \sqrt{c_t}\Pi_1 1_K \\ \vdots \\ \sqrt{c_t}\Pi_t 1_K \end{bmatrix} = 0 \tag{B.1}$$

almost surely. Assume that, for all t, $\mathbb{1}(T > t+1)$ *and* Π_{t+1} *are measurable with respect to* $B_1 \mathbf{Y}_{1:1}, \ldots, B_t \mathbf{Y}_{1:t}$*. Consider* $\sqrt{n}(L - \lambda'\mu) = 1_K' \Sigma^{-1} \sum_{t=1}^{T-1} \sqrt{c_t} Y_t$*. Then*

$$\begin{bmatrix} \sqrt{n}(L - \lambda'\mu) \\ Y_T \end{bmatrix} \mid T, B_1 \mathbf{Y}_{1:1}, \ldots, B_{T-1}\mathbf{Y}_{1:T-1} \sim \mathcal{N}\left(0, \begin{bmatrix} 1'\left(\Sigma^{-1} \sum_{t=1}^{T-1} c_t \Pi_t\right) 1 & 0 \\ 0 & \Pi_T \Sigma \end{bmatrix}\right).$$

Proof. We show the following claim with a fixed stopping time: For all $s = 2, \ldots, T_0$, let

$$Z_s = 1_K' \Sigma^{-1} \sum_{t=1}^{s-1} \sqrt{c_t} Y_t.$$

Then, for all s,

$$(Z_s, Y_s) \mid B_1 \mathbf{Y}_{1:1}, \ldots, B_{s-1} \mathbf{Y}_{1:s-1} \sim \mathcal{N}\left(0, \operatorname{diag}\left(1'\left(\Sigma^{-1} \sum_{t=1}^{s-1} c_t \Pi_t\right) 1, \Pi_s \Sigma\right)\right). \tag{B.2}$$

Given (B.2), note that for all s,

$$\{(Z_T, Y_T) \mid T = s, B_1 \mathbf{Y}_{1:1}, \ldots, B_{T-1}\mathbf{Y}_{1:T-1}\} \sim \{(Z_s, Y_s) \mid B_1 \mathbf{Y}_{1:1}, \ldots, B_{s-1} \mathbf{Y}_{1:s-1}\},$$

by our assumptions about the stopping time T. This proves this lemma.

To show (B.2), we induct on t. To that end, define

$$\lambda_t = \begin{bmatrix} \sqrt{c_t}\Sigma^{-1} 1_K \\ \vdots \\ \sqrt{c_t}\Sigma^{-1} 1_K \end{bmatrix} \equiv \begin{bmatrix} \tilde{\lambda}_1 \\ \vdots \\ \tilde{\lambda}_t \end{bmatrix}$$

We wish to verify the following claims for all $t = 1, \ldots, T-1$:

(1) $\mathbf{Y}_{1:t+1} \mid \{B_s \mathbf{Y}_{1:s} : s \le t\} \sim \mathcal{N}\left(\mu_{t+1|t}, \Sigma_{t+1|t}\right)$ where

$$\Sigma_{t+1|t} = \begin{bmatrix} \Sigma_{t|t} & 0 \\ 0 & \Sigma\Pi_{t+1} \end{bmatrix} \quad \Sigma_{t|t} = \Sigma_{t|t-1} - \Sigma_{t|t-1} B_t' (B_t \Sigma_{t|t-1} B_t')^+ B_t \Sigma_{t|t-1},$$

$$\mu_{t+1|t} = \begin{bmatrix} \mu_{t|t} \\ 0 \end{bmatrix} \quad \mu_{t|t} = \mu_{t|t-1} + \Sigma_{t|t-1} B_t' (B_t \Sigma_{t|t-1} B_t')^+ B_t (\mathbf{Y}_{1:t} - \mu_{t|t-1})$$

$$\Sigma_{1|0} = \Sigma\Pi_1$$

$$\mu_{1|0} = 0$$

(2)

$$\Sigma_{t|t}\lambda_t = \Sigma_{t|t-1}\lambda_t = \begin{bmatrix} \sqrt{c_1}\Pi_1 1_K \\ \vdots \\ \sqrt{c_t}\Pi_t 1_K \end{bmatrix}$$

(3) $B_t \Sigma_{t|t-1}\lambda_t = 0$

(4) $\lambda_t' \mu_{t|t} = 0$

Assuming items (1)–(4), we show (B.2). Note that $Z_s = \lambda_{s-1}' \mathbf{Y}_{1:s-1}$. Hence,

$$\begin{bmatrix} Z_s \\ Y_s \end{bmatrix} = \begin{bmatrix} \lambda_{s-1}' & 0 \\ 0 & I_K \end{bmatrix} \mathbf{Y}_{1:s}$$

and is jointly Gaussian. Its mean conditional on $\{B_t \mathbf{Y}_{1:t} : t \le s-1\}$ is thus

$$\begin{bmatrix} \lambda_{s-1}' & 0 \\ 0 & I_K \end{bmatrix} \begin{bmatrix} \mu_{s-1|s-1} \\ 0 \end{bmatrix} = 0$$

and its conditional variance is

$$\begin{aligned} &\begin{bmatrix} \lambda_{s-1}' & 0 \\ 0 & I_K \end{bmatrix} \begin{bmatrix} \Sigma_{s-1|s-1} & 0 \\ 0 & \Sigma\Pi_s \end{bmatrix} \begin{bmatrix} \lambda_{s-1}' & 0 \\ 0 & I_K \end{bmatrix}' \\ &= \begin{bmatrix} \lambda_{s-1}'\Sigma_{s-1|s-1}\lambda_{s-1} & 0 \\ 0 & \Sigma\Pi_s \end{bmatrix} = \begin{bmatrix} 1_K' \left(\Sigma^{-1}\sum_{t=1}^{s-1} c_t \Pi_t\right) 1_K & 0 \\ 0 & \Sigma\Pi_s \end{bmatrix} \end{aligned}$$

as desired. Thus, items (1)–(4) prove (B.2).

Finally, to verify items (1)–(4), consider the base case $t = 1$. Then $\mathbf{Y}_1 = Y_1 \sim \mathcal{N}(0, \Sigma\Pi_1) = \mathcal{N}(\mu_{1|0}, \Sigma_{1|0})$. Note that

$$Y_1 \mid B_1 Y_1 \sim \mathcal{N}\left(\underbrace{\mu_{1|0} - \Sigma_{1|0} B_1' (B_1 \Sigma_{1|0} B_1')^+ B_1 (Y_1 - \mu_{1|0})}_{\mu_{1|1}}, \underbrace{\Sigma_{1|0} - \Sigma_{1|0} B_1' (B_1 \Sigma_{1|0} B_1')^+ B_1 \Sigma_{1|0}}_{\Sigma_{1|1}} \right)$$

and $(Y_2 \mid B_1 Y_1) \sim (Y_2 \mid \Pi_2) \sim \mathcal{N}(0, \Pi_2 \Sigma)$. Thus $\mathbf{Y}_{1:2} \mid B_1 \mathbf{Y}_{1:1} \sim \mathcal{N}(\mu_{2|1}, \Sigma_{2|1})$. This verifies item (1). For items (2) and (3), note that $\Sigma_{1|0}\lambda_1 = \Sigma\Pi_1\sqrt{c_t}\Sigma^{-1}1_K = \sqrt{c_t}\Pi_1 1_K$. Note that $B_1\Sigma_{1|0}\lambda_1 = B_1\sqrt{c_t}\Pi_1 1_K = 0$ by assumption. Hence $\Sigma_{1|1}\lambda_1 = \Sigma_{1|0}\lambda_1$. This verifies items (2) and (3). Finally, $\lambda_1 \mu_{1|1} = \lambda' \mu_{1|0} = 0$ by (3). This verifies item (4).

For the inductive case, assume that items (1)–(4) holds for all $s \le t-1$. Then

$$\mathbf{Y}_{1:t} \mid \{B_s \mathbf{Y}_{1:s} : s \le t-1\} \sim \mathcal{N}(\mu_{t|t-1}, \Sigma_{t|t-1})$$

is jointly Gaussian. Hence

$$\mathbf{Y}_{1:t} \mid \{B_s \mathbf{Y}_{1:s} : s \le t\} \sim \mathcal{N}\left(\mu_{t|t}, \Sigma_{t|t}\right).$$

Note too that

$$Y_{t+1} \mid \{B_s \mathbf{Y}_{1:s} : s \le t\} \sim (Y_{t+1} \mid \Pi_{t+1}) \sim \mathcal{N}\left(0, \Pi_{t+1}\Sigma\right).$$

independently from $\mathbf{Y}_{1:t}$. This verifies item (1) for t.

For items (2) and (3), note that

$$\Sigma_{t|t-1}\lambda_t = \begin{bmatrix} \Sigma_{t-1|t-1}\lambda_{t-1} \\ \Sigma\Pi_t\tilde{\lambda}_t \end{bmatrix} = \begin{bmatrix} \sqrt{c_1}\Pi_1 1_K \\ \vdots \\ \sqrt{c_{t-1}}\Pi_{t-1}1_K \\ \sqrt{c_t}\Pi_t 1_K \end{bmatrix} \text{ and } B_t\Sigma_{t|t-1}\lambda_t = B_t \begin{bmatrix} \sqrt{c_1}\Pi_1 1_K \\ \vdots \\ \sqrt{c_t}\Pi_t 1_K \end{bmatrix} = 0$$

by assumption. Hence $\Sigma_{t|t-1}\lambda_t = \Sigma_{t|t}\lambda_t$. Finally, for (4), $\lambda_t'\mu_{t|t} = \lambda_t'\mu_{t|t-1} = \lambda_{t-1}'\mu_{t-1|t-1} = 0$ by (3). This completes the proof. □

B.3. **Sufficiency in Thompson sampling.**

Proposition 3.1. *Under the setup of Theorem 3.7, suppose, for all t, $\Xi_t(X_1, \ldots, X_{t-1})$ is measurable with respect to $\Pi_{2:t}$. Under Thompson sampling, (L, X_T) is sufficient for μ with respect to the conditional distribution $X_{1:T} \mid \Pi_{1:T}, T$.*

Proof. We will show the following two claims:

(1) Under Thompson sampling, $(L, \Pi_{2:T})$ is sufficient for μ with respect to $X_{1:T} \mid T$.
    This follows directly from Lemma B.2.
(2) If $(L, \Pi_{2:T})$ is sufficient for μ with respect to $X_{1:T} \mid T$, then (L, X_T) is sufficient for μ with respect to $X_{1:T} \mid T, \Pi_{2:T}$.
    Observe that

$$\begin{aligned} p_\mu(X_{1:T} \mid T) &= p_\mu(X_{1:T-1} \mid T)p_\mu(X_T \mid X_{1:T-1}, T) \\ &= g(X_{1:T-1} \mid \Pi_{2:T}, L, T)p_\mu(\Pi_{2:T}, L \mid T)p_\mu(X_T \mid \Pi_{2:T}, T) \end{aligned}$$

where g does not depend on μ since $(\Pi_{2:T}, L)$ is sufficient with respect to $X_{1:T-1} \mid T$. Since μ enters only through functions that depend on $L, X_T, \Pi_{2:T}$, we conclude that $(L, X_T, \Pi_{2:T})$ is sufficient for μ with respect to $X_{1:T} \mid T$. Thus, (L, X_T) is sufficient for μ with respect to $X_{1:T} \mid \Pi_{2:T}, T$.

□

Lemma B.2. *Under Thompson sampling, assume that Ξ_t is measurable with respect to $\Pi_{2:t}$. Then $(\Pi_{2:T}, L)$ is sufficient for μ with respect to $X_{1:T-1} \mid T$.*

Proof. Let $L_s = \sum_{k=1}^K \sum_{t=1}^{s-1} \frac{n_t \Pi_{tk}}{n\sigma_k^2} X_{tk}$. We shall show the claim that $(\Pi_{2:s}, L_s)$ is sufficient for μ with respect to $X_{1:s-1}$. Given this result (which has a fixed stopping time), we note that, since $\{T = s\}$ is measurable with respect to $\Pi_{2:s}$,

$$p(X_{1:T-1} \mid T = s, \Pi_{2:T}, L) = p(X_{1:s-1} \mid T = s, \Pi_{2:s}, L_s) = p(X_{1:s-1} \mid \Pi_{2:s}, L_s).$$

Since $(\Pi_{2:s}, L_s)$ is sufficient for μ with respect to $X_{1:s-1}$, the above display does not depend on μ. As a result, $(\Pi_{2:T}, L)$ is sufficient for μ with respect to $X_{1:T-1} \mid T$. This proves Lemma B.2 provided the claim holds.

To prove the claim, note that by the Hotz and Miller (1993) inversion result (Norets and Takahashi, 2013), Q is a bijection between Π_{t+1} and the differences

$$\nabla_t = (W_{t,1} - W_{t,K}, \ldots, W_{t,K-1} - W_{t,K}).$$

Thus, it suffices to show that $(\nabla_{1:s-1}, L_s)$ is sufficient for μ with respect to $X_{1:s-1}$.

Note that

$$\begin{aligned}
L_s &= \sum_{k=1}^{K} \sum_{t=1}^{s-1} \frac{n_t \Pi_{tk}}{n\sigma_k^2} X_{tk} \\
&= \sum_{k=1}^{K} \left(\sum_{t=1}^{s-1} \frac{n_t \Pi_{tk}}{n\sigma_k^2} \right) W_{s-1,k} \\
L_s - \sum_{k=1}^{K-1} \left(\sum_{t=1}^{s-1} \frac{n_t \Pi_{tk}}{n\sigma_k^2} \right) (W_{s-1,k} - W_{s-1,K}) &= \left(\sum_{s=1}^{s-1} \frac{n_t \Pi_{tK}}{n\sigma_K^2} \right) W_{s-1,K} + \sum_{k=1}^{K-1} \left(\sum_{t=1}^{s-1} \frac{n_t \Pi_{tk}}{n\sigma_k^2} \right) W_{s-1,K} \\
&= W_{s-1,K} \sum_{t=1}^{s-1} \sum_{k=1}^{K} \frac{n_t \Pi_{tk}}{n\sigma_k^2}.
\end{aligned}$$

That is, we can transform $(\nabla_{1:s-1}, L_s)$ into $(\Pi_{1:s}, W_{s-1})$, the latter of which is sufficient for μ with respect to $X_{1:s-1}$. This completes the proof. $\square$

B.4. **Polyhedral assignment.**

Theorem 4.2. *Assume, without loss of generality, that $\|\eta\| = 1$. Define $R = n \sum_{t=1}^{T} V_t^{-1}$. Let $G = 1_T' \otimes I_K = [I_K, \ldots, I_K]$ be a $K \times TK$ matrix. Let $G_\perp$ be a $(T-1)K \times TK$ matrix whose rows are unit vectors that are orthogonal to each other and to the rows of G. Let $\mathbf{V} = \frac{1}{n} \operatorname{diag}(V_1, \ldots, V_T)$ be a $TK \times TK$ diagonal matrix. Let $\tilde{G}, \tilde{G}_\perp$ partition the inverse of the following matrix:*

$$\begin{bmatrix} G\mathbf{V}^{-1} \\ G_\perp \end{bmatrix}^{-1} = \left[\tilde{G}, \tilde{G}_\perp \right].$$

where $\tilde{G}$ is $TK \times K$. Let $c = \frac{1}{\eta' R^{-1} \eta}$. Then

$$\left(\frac{U}{c} \mid U_\perp = u_\perp, A(\mathcal{F}_{T-1}) \mathbf{X}_{1:T} \leq b(\mathcal{F}_{T-1}) \right) \sim (Z_1 \mid M(\mathcal{F}_{T-1}) Z \leq m(\mathcal{F}_{T-1}))$$

where the right-hand side follows the law

$$Z = (Z_1, Z_2')' \sim \mathcal{N} \left(\begin{bmatrix} \tau + k' u_\perp \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{c} & 0 \\ 0 & G_\perp \mathbf{V} G_\perp' \end{bmatrix} \right) \qquad k' = \frac{1}{c} \eta' G \mathbf{V}^{-1} G' \eta_\perp' (\eta_\perp G \mathbf{V}^{-1} G' \eta_\perp')^{-1}$$

The constraints are defined by

$$M(\mathcal{F}_{T-1}) = \left[cA(\mathcal{F}_{T-1})\tilde{G}\eta, A(\mathcal{F}_{T-1})\tilde{G}_\perp\right] \quad m(\mathcal{F}_{T-1}) = b(\mathcal{F}_{T-1}) - A(\mathcal{F}_{T-1})\tilde{G}(I - \eta\eta')S.$$

Proof. Note that $S = G\mathbf{V}^{-1}\mathbf{X}_{1:T}$ and that $R = G\mathbf{V}^{-1}G'$. We are interested in the joint distribution of

$$\begin{aligned} \begin{bmatrix} U \\ U_\perp \end{bmatrix} \mid \Pi_{2:T} = \pi &\sim \begin{bmatrix} U \\ U_\perp \end{bmatrix} \mid A(\mathcal{F}_{T-1})\mathbf{X}_{1:T} \leq b(\mathcal{F}_{T-1}) \\ &\sim \begin{bmatrix} \eta' G\mathbf{V}^{-1} \\ \eta'_\perp G\mathbf{V}^{-1} \end{bmatrix} \mathbf{X}_{1:T} \mid A(\mathcal{F}_{T-1})\mathbf{X}_{1:T} \leq b(\mathcal{F}_{T-1}) \\ &\sim \begin{bmatrix} \eta' G\mathbf{V}^{-1} \\ \eta'_\perp G\mathbf{V}^{-1} \end{bmatrix} \mathbf{X}^*_{1:T} \mid A(\mathcal{F}_{T-1})\mathbf{X}^*_{1:T} \leq b(\mathcal{F}_{T-1}) \end{aligned}$$

where $\mathbf{X}^*_{1:T} \sim \mathcal{N}(1_T \otimes \mu, \mathbf{V})$. Let $S^* = G\mathbf{V}^{-1}\mathbf{X}^*_{1:T}$. Thus, it suffices to study the conditional distribution

$$\underbrace{\eta' G\mathbf{V}^{-1}\mathbf{X}^*_{1:T}}_{cZ_1} \mid \underbrace{\eta'_\perp G\mathbf{V}^{-1}\mathbf{X}^*_{1:T}}_{Z_3}, A(\mathcal{F}_{T-1})\mathbf{X}^*_{1:T} \leq b(\mathcal{F}_{T-1}).$$

Let $G_\perp \mathbf{X}^*_{1:T} = Z_2$.

We first study the Gaussian distribution $(Z_1, Z_2') \mid Z_3$. This amounts to calculating conditional means and variances. Note that

$$\begin{aligned} \mathbb{E}[cZ_1 \mid Z_3] &= \eta' R\mu + ck'(Z_3 - \eta_\perp R\mu) \\ \eta' R\mu - ck'\eta_\perp R\mu &= \eta' R\mu - \eta' R\eta'_\perp(\eta_\perp R\eta'_\perp)^{-1}\eta_\perp R\mu \\ &= \eta' R\mu - \eta' R\eta'_\perp(\eta_\perp R\eta'_\perp)^{-1}\eta_\perp R(\eta\eta' + \eta'_\perp\eta_\perp)\mu \\ &= \eta' R\mu - \left(\eta' R\eta'_\perp(\eta_\perp R\eta'_\perp)^{-1}\eta_\perp R\eta\right) \cdot \tau - \eta' R\eta'_\perp\eta_\perp\mu \\ &= \eta' R(I - \eta'_\perp\eta_\perp)\mu - \left(\eta' R\eta'_\perp(\eta_\perp R\eta'_\perp)^{-1}\eta_\perp R\eta\right) \cdot \tau \\ &= \left(\eta' R\eta - \eta' R\eta'_\perp(\eta_\perp R\eta'_\perp)^{-1}\eta_\perp R\eta\right) \tau \end{aligned}$$

and hence

$$\mathbb{E}[cZ_1 \mid Z_3] = ck'Z_3 + \left(\eta' R\eta - \eta' R\eta'_\perp(\eta_\perp R\eta'_\perp)^{-1}\eta_\perp R\eta\right) \tau.$$

Next, we show that

$$c = \eta' R\eta - \eta' R\eta'_\perp(\eta_\perp R\eta'_\perp)^{-1}\eta_\perp R\eta = \eta' R^{1/2}\left(I - R^{1/2}\eta'_\perp(\eta_\perp R\eta'_\perp)^{-1}\eta_\perp R^{1/2}\right) R^{1/2}\eta$$

The matrix $\left(I - R^{1/2}\eta'_\perp(\eta_\perp R\eta'_\perp)^{-1}\eta_\perp R^{1/2}\right)$ projects to the space of all vectors perpendicular to $\eta_\perp R^{1/2}$. Note that this space is one dimensional with basis vector $R^{-1/2}\eta$. As a result,

$$I - R^{1/2}\eta'_\perp(\eta_\perp R\eta'_\perp)^{-1}\eta_\perp R^{1/2} = \frac{R^{-1/2}\eta\eta' R^{-1/2}}{\eta' R^{-1}\eta}.$$

Hence

$$\eta' R^{1/2} \left(I - R^{1/2}\eta_\perp'(\eta_\perp R\eta_\perp')^{-1}\eta_\perp R^{1/2}\right) R^{1/2}\eta = \frac{1}{\eta' R^{-1}\eta} = c.$$

Hence

$$\mathbb{E}[Z_1 \mid Z_3] = \tau + k' Z_3.$$

Note that $\mathrm{Cov}(S^*, Z_2) = G\mathbf{V}^{-1}\mathbf{V}G_\perp' = 0$ by construction. Hence $\mathrm{Cov}(Z_1, Z_2 \mid Z_3) = 0$, $\mathrm{Var}(Z_2 \mid Z_3) = G_\perp \mathbf{V} G_\perp'$, and $\mathbb{E}[Z_2 \mid Z_3] = 0$. Lastly,

$$\mathrm{Var}(cZ_1 \mid Z_3) = \eta' R\eta - \eta' R\eta_\perp'(\eta_\perp R\eta_\perp')^{-1}\eta' R\eta = c$$

and hence

$$\mathrm{Var}(Z_1 \mid Z_3) = 1/c.$$

This implies that

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \mid Z_3 \sim \mathcal{N}\left(\begin{bmatrix} \tau + k' Z_3 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/c & 0 \\ 0 & G_\perp \mathbf{V} G_\perp' \end{bmatrix}\right).$$

It remains to show that the constraint $A(\mathcal{F}_{T-1})\mathbf{X}_{1:T}^* \leq b(\mathcal{F}_{T-1})$ is equivalent to $M(\mathcal{F}_{T-1})Z \leq b(\mathcal{F}_{T-1})$. This is true by plugging in

$$\mathbf{X}_{1:T}^* = \tilde{G}S^* + \tilde{G}_\perp Z_2 = \tilde{G}(cZ_1) + \tilde{G}\underbrace{(I - \eta\eta')}_{\eta_\perp'\eta_\perp} S^* + \tilde{G}_\perp Z_2.$$

This completes the proof. □

Appendix C. Proofs for results in Appendix A

C.1. **Outline of the general argument.** We recall the following notation choices in Appendix A and define some more notation.

- Recall that there are at most $T_0 > 0$ batches with an adaptive stopping time T_n. Assume that the sample sizes $n_t/n \to c_t$ for all $t = 1, \ldots, T_0$ are fixed sequences.
- Let Π_{tkn} be the algorithm-determined probability of assigning to arm k in batch t
- Let $N_{tkn} \mid \Pi_{tkn} \sim \mathrm{Bin}(n_t, \Pi_{tkn})$ be the realized number of samples assigned to arm k in batch t, where $N_{tkn} = \sum_{i \in \mathcal{I}_t} D_{ik}$.
- Let $\hat{\Pi}_{tkn} = N_{tkn}/n_t$.
- Let $X_{tkn} = \frac{1}{N_{tkn} \vee 1} \sum_{i \in \mathcal{I}_t} D_{ik} X_i^{\mathrm{obs}}$ be the mean arm value for batch t and arm k. If no one is assigned to arm k, then $X_{tkn} = 0$. Let X_{tn} collect X_{tkn}.
- Let $Y_{tkn} = \frac{N_{tkn}}{\sqrt{n_t}}(X_{tkn} - \mu_{nk}) = \sqrt{n_t}\hat{\Pi}_{tkn}(X_{tkn} - \mu_{nk}) = \frac{1}{\sqrt{n_t}} \sum_{i \in \mathcal{I}_t} D_{ik}(X_i^{\mathrm{obs}} - \mu_{nk})$ be a scaled version of X_{tkn}. Let Y_{tn} collect Y_{tkn}.
- Let $Y_{:t,k,n} = \sum_{s=1}^t \sqrt{n_s/n} Y_{s,k,n}$
- Let $\hat{\Pi}_{:t,k,n} = \sum_{s=1}^t \hat{\Pi}_{s,k,n} \frac{n_s}{n}$

- Let
$$W_{t,k,n} = \frac{\sum_{s=1}^{t} N_{skn} X_{skn}}{\sum_{s=1}^{t} N_{skn}}$$
be the cumulative arm-k mean where
$$\sqrt{n}(W_{t,k,n} - \mu_{nk}) = \frac{1}{\hat{\Pi}_{:t,k,n}} Y_{:t,k,n}.$$
- Let (A.1) be an estimate of the arm variance using data up until batch t. Let $\hat{\sigma}^2_{kn} = \hat{\sigma}^2_{Tkn}$.
- Let
$$L_n = \sum_{t=1}^{T_n-1} \sum_{k=1}^{K} \frac{n_t \hat{\Pi}_{tkn}}{\sqrt{n}\hat{\sigma}^2_{kn}} X_{tkn}$$
be the properly scaled empirical analogue to L.
- Let $\lambda_{kn} = \sum_{t=1}^{T-1} \frac{n_t \hat{\Pi}_{tkn}}{\sqrt{n}\hat{\sigma}^2_{kn}}$ and $\lambda_n = [\lambda_{1n}, \ldots, \lambda_{Kn}]'$. Note that
$$L_n - \lambda_n' \mu_n = \sum_{t=1}^{T_n-1} \sum_{k=1}^{K} \frac{n_t \hat{\Pi}_{tkn}}{\sqrt{n}\hat{\sigma}^2_{kn}} (X_{tkn} - \mu_{nk}) = \sum_{t=1}^{T_n-1} \sum_{k-1}^{K} \frac{\sqrt{n_t/n}}{\hat{\sigma}^2_{kn}} Y_{tkn}.$$

Our goal is to show that (A.2) converges to a Gaussian random variable, and its limit is conditionally Gaussian given limits of $\Pi_{1:T,n}, T_n, \eta_n$. To do so, we first analyze the behavior of $(\Pi_{1:T_0,n}, Y_{1:T_0,n})$ without a stopping time—we could imagine that the experiment carries on after T_n. Using the scaling Y instead of X allows for Π_{tkn} to be very close to zero or actually zero.

We show $(\Pi_{1:T_0,n}, Y_{1:T_0,n})$ weakly converges by the following iterative process. So long as Π_{t+1} is a suitably continuous function of the past information $G_{t,n} = (\Pi_{1:t,n}, Y_{1:t,n})$, it weakly converges to a function of the limits of the past information $\Pi_{t+1}(G_t)$. Given this convergence, we also show that Y_{t+1} converges to $\mathcal{N}(0, \Pi_{t+1}\Sigma)$, and hence $G_{t+1,n}$ converges (Theorem C.1).

Second, given a stopping time T_n that is a suitably continuous function of the past information, we can further show that $L_n, Y_{T_n,n}$ converges to quantities that behave like their Gaussian experiment counterparts (Corollary C.1).

Third, given suitably continuous η_n, we can show that the statistic (A.2), properly normalized, converges to a random variable Z where $Z \mid \Pi_{1:T}, T, \eta \sim \mathcal{N}(0,1)$, where $\Pi_{1:T}, T, \eta$ are limits of their finite-sample analogues (Theorem C.2). Since Z is conditionally Gaussian, for sufficiently continuous weighting functions f, we can show that f-weighted average mis-coverage rates converge to zero uniformly over $\mathcal{P}$ (Theorem C.3).

Finally, we verify Theorem A.2 by verifying that the continuity restrictions on $\Pi_t, \mathbb{1}(T_n > t), \eta_n$ are satisfied given the assumptions in Appendix A.

The uniformity aspects of the argument use the subsequencing argument in Andrews et al. (2011). The key to doing so is the following lemma, which establishes that certain subsequences exist given our assumptions on $\mathcal{P}$.

Lemma C.1. *Under Assumption A.1, for any sequence $P_n \in \mathcal{P}$ and any subsequence n_r of n, there exists a further subsequence n_p of n_r such that, as $p \to \infty$,*

(1) $\mu_{n_p} \equiv \mu(P_{n_p}) \to \mu \in [-\infty, \infty]^K$,

(2) $h_{n_p} \equiv \sqrt{n_p}(\mu_{n_p} - \max_k \mu_{n_p,k}) \to h \in [-\infty, 0]^K$,

(3) For all $k \neq k' \in [K]$, $\Delta\mu_{k,k',n_p} \equiv \sqrt{n}(\mu_{k,n_p} - \mu_{k',n_p}) \to \Delta\mu_{k,k'} \in [-\infty, \infty]$,

(4) For all k, $\sigma_k^2(P_{n_p}) \to \sigma_k^2 \in [C_1^{-1}, C_1]$,

(5) For all $k \in [K], t \in [T_0]$, $\hat{\sigma}^2_{tkn_p} \xrightarrow{p} \sigma_k^2$.

Proof. Fix the subsequence n_r and the sequence P_n. There exists a subsequence for which (1)–(4) are satisfied since $[-\infty, \infty]$, $[-\infty, 0]$ (under the metric $d(x, y) = |\arctan(x) - \arctan(y)|$), and $[C_1^{-1}, C_1]^K$ are all compact.

Lastly, for (5), note that (A.1) can be written as

$$\sigma^2_{tkn} = -(W_{tkn} - \mu_k(P_n))^2 + \frac{1}{\sum_{s=1}^t N_{skn}} \sum_{i \in \mathcal{I}_{1:t}} D_{ik}(X_i^{\text{obs}} - \mu_k(P_n))^2.$$

We verify the following claims:

(a) We have

$$\mathrm{P}\left[\left|\sum_{i \in \mathcal{I}_{1:t}} D_{ik}(X_i^{\text{obs}} - \mu_k(P_n))\right| > \sqrt{n \log n}\right] \leq \frac{C_1}{\log n}.$$

(b) We have

$$\frac{1}{\sum_{s=1}^t N_{skn}} = O_p(1/n)$$

(c) For some $C > 0$,

$$\mathrm{P}\left[\left|\sum_{i \in \mathcal{I}_{1:t}} D_{ik}((X_i^{\text{obs}} - \mu_k(P_n))^2 - \sigma_k^2(P_n))\right| > \sqrt{n \log n}\right] \leq \frac{C}{\log n}.$$

Given these claims, note that (a) and (b) imply that $W_{tkn} - \mu_k(P_n) = o_p\left(\sqrt{\frac{\log n}{n}}\right)$ and (b) and (c) imply that

$$\frac{1}{\sum_{s=1}^t N_{skn}} \sum_{i \in \mathcal{I}_{1:t}} D_{ik}(X_i^{\text{obs}} - \mu_k(P_n))^2 - \sigma_k^2(P_n) = o_p\left(\sqrt{\frac{\log n}{n}}\right).$$

As a result, $\sigma^2_{tkn} - \sigma_k^2(P_n) = o_p(\sqrt{\log n/n})$. Hence $\sigma^2_{tkn} - \sigma_k^2 = o_p(1)$.

For (a), we note that the sum is a martingale. By Kolmogorov's inequality for martingales,

$$\mathrm{P}\left(\left|\sum_{i\in\mathcal{I}_{1:t}} D_{ik}(X_i^{\mathrm{obs}}-\mu_k(P_n))\right| > \sqrt{n\log n}\right) \le \frac{1}{n\log n}\sum_{i=1}^n \sigma_k^2(P_n) \le \frac{C_1}{\log n}.$$

For (b), we note that $\sum_s N_{skn} > N_{1kn} \sim \mathrm{Bin}(c_1 n, \pi_{1k})$. Since N_{1kn} grows at the rate of n, $\frac{1}{\sum_{s=1}^t N_{skn}} = O_p(1/n)$. Similarly, we recognize that the sum in (c) is again a martingale. To apply Kolmogorov's inequality, we need that the second moment of $(X_i^{\mathrm{obs}}-\mu_k(P_n))^2-\sigma_k^2(P_n)$ is finite. This is true due to the fourth moment condition in Assumption A.1. $\square$

C.2. **Behavior of random variables without a stopping time.** We first analyze the limiting behavior of the various random variables without a stopping time. That is, we consider letting the experiment run until T_0.

Recall that $\Pi_{tn} = \mathrm{diag}(\Pi_{t1n},\ldots,\Pi_{tKn})$ collect the batch t probabilities and that $\hat\Sigma_{tn}$ collect $\hat\sigma^2_{tkn}$. We assume that, for some function $\kappa_{t+1,n}$,

$$\Pi_{t+1,n} = \kappa_{t+1,n}\left(X_{1:t,n},\hat\Pi_{1:t,n},\hat\Sigma_{tn},n_{1:t}\right).$$

Note that

$$X_{tkn} = \begin{cases}\frac{1}{\sqrt{n_t\hat\Pi_{tkn}}}Y_{tkn}+\mu_{nk}, & \text{if } \hat\Pi_{tkn}>0\\ 0 & \text{otherwise.}\end{cases}$$

and thus is a function of $n_t, \hat\Pi_{tn}, Y_{tn}$. Therefore, we can write instead

$$\Pi_{t+1,n} = \kappa_{t+1,n}\left(Y_{1:t,n},\hat\Pi_{1:t,n},\hat\Sigma_{tn},n_{1:t},\mu_n\right). \tag{C.1}$$

We require that κ_{t+1} satisfies the following implication of the continuous mapping theorem. This is more general than the conditions in Appendix A.

Assumption C.1. *The assignment algorithm $(\pi_1,\kappa_2,\ldots,\kappa_{T_0})$, for κ_t in (C.1), satisfies the following properties: Fix some $\epsilon>0$,*

(1) The assignment probabilities in the first batch are fixed and bounded away from zero: $\pi_{1k}>\epsilon>0$ for every $k\in[K]$. Let $\Pi_1=\pi_1$.

(2) Let P_{n_m} denote a (sub)sequence of data-generating processes satisfying the conclusion of Lemma C.1. Let $\Sigma=\mathrm{diag}(\sigma_1^2,\ldots,\sigma_K^2)$. Assume that, if

$$Y_{1n_m}\xrightarrow{d} Y_1\sim\mathcal{N}(0,\Pi_1\Sigma)\quad \hat\Pi_{1n_m}\xrightarrow{p}\Pi_1, \tag{C.2}$$

then for some measurable function κ_2, we have the following convergence jointly with (C.2)

$$\Pi_{2n_m} = \kappa_{2,n_m}\left(Y_{1,n_m},\hat\Pi_1,\hat\Sigma_{1n_m},n_1,\mu_{n_m}\right)\xrightarrow{d}\kappa_2\left(Y_1,\Pi_1;h,\sigma^2_{1:K},c_1\right)\equiv\Pi_2.$$

Moreover, assume that, if for some given $1<t<T_0$,

$$(Y_{1n_m},\Pi_{2n_m},\ldots,\Pi_{tn_m},Y_{tn_m})\xrightarrow{d}(Y_1,\Pi_2,\ldots,\Pi_t,Y_t)\quad \textit{and}\quad \hat\Pi_{1:t,n_m}\xrightarrow{p}\Pi_{1:t}, \tag{C.3}$$

where for all s, $Y_s \mid Y_{1:s-1}, \Pi_{2:s} \sim \mathcal{N}(0, \Pi_t \Sigma)$, then there is a measurable function κ_{t+1} such that we have the following convergence jointly with (C.3)

$$\Pi_{t+1,n} = \kappa_{t+1,n}\left(Y_{1:t,n}, \hat{\Pi}_{1:t,n}, \hat{\Sigma}_{tn}, n_{1:t}, \mu_n\right) \xrightarrow{d} \kappa_{t+1}\left(Y_{1:t}, \Pi_{1:t}; h, \sigma^2_{1:K}, c_{1:t}\right) \equiv \Pi_{t+1}. \quad \text{(C.4)}$$

(3) For all $t = 2, \ldots, T_0$, the limiting random variable Π_t is measurable with respect to

$$\left(B_{t-1}(\Pi_{1:t-1}, c_{1:t-1}) \begin{bmatrix} Y_1 \\ \vdots \\ Y_{t-1} \end{bmatrix}, \Pi_{1:t-1} \right)$$

for some conformable matrix $B_{t-1}(\Pi_{1:t-1}, c_{1:t-1})$ such that

$$B_{t-1}(\Pi_{1:t-1}, c_{1:t-1}) \begin{bmatrix} \sqrt{c_1}\Pi_1 1_K \\ \vdots \\ \sqrt{c_{t-1}}\Pi_{t-1} 1_K \end{bmatrix} = 0$$

almost surely.

Let $G_{tn} = (\Pi_{1:t,n}, Y_{1:t,n})$ and let $G_t = (\Pi_{1:t}, Y_{1:t})$. The following theorem characterizes the asymptotic behavior of $G_{T_0,n}$ and connects to the normal model results.

Theorem C.1. *Under Assumption A.1, Assumption C.1(1), and Assumption C.1(2), for any sequence $P_n \in \mathcal{P}$ and any subsequence n_r of n, there exists a further subsequence n_m such that, as $m \to \infty$,*

$$G_{T_0,n_m} \xrightarrow{d} G_{T_0} = (\Pi_{1:T_0}, Y_{1:T_0})$$

where, for all $t \in [T_0]$,

$$Y_t \mid Y_{1:t-1}, \Pi_{1:t} \sim \mathcal{N}(0, \Pi_t \Sigma).$$

and Π_{t+1} is measurable with respect to $\{B_s \mathbf{Y}_{1:s} : s \leq t\}$.

Proof. Fix P_n and a subsequence n_r of n. By Lemma C.1, there exists a further subsequence n_m such that the conclusions of Lemma C.1 hold.

We induct on t. Consider the base case of $G_{1n} = (\Pi_{1n}, Y_{1n})$. By assumption Π_{1n} is fixed and constant in n. Thus it suffices to show that

$$Y_{1n_m} \xrightarrow{d} \mathcal{N}(0, \Pi_1 \Sigma).$$

This is true by the moment condition in Assumption A.1, where we apply the central limit theorem in Lemma C.2.

For the inductive case, suppose $G_{tn_m} \xrightarrow{d} (\Pi_{1:t}, Y_{1:t})$ for some $1 \leq t < T_0$. We wish to show that $G_{t+1,n_m} \xrightarrow{d} (\Pi_{1:t+1}, Y_{1:t+1})$. First, note that by a simple Chernoff bound in Lemma C.3, for all $s \leq t$,

$$\hat{\Pi}_{s,n} - \Pi_{s,n} = o_p(1).$$

Hence, $\hat{\Pi}_{s,n} \xrightarrow{p} \Pi_s$ under the inductive hypothesis. Now, by our assumptions on the assignment algorithm in Assumption C.1(2),

$$(\Pi_{1:t+1,n_m}, Y_{1:t,n_m}) \xrightarrow{d} (\Pi_{1:t+1}, Y_{1:t}). \tag{C.5}$$

By Assumption C.1(2), Π_{t+1} is measurable with respect to $(B_t \mathbf{Y}_{1:t}, \Pi_{1:t})$, which is further measurable with respect to $\{B_s \mathbf{Y}_{1:s} : s \le t\}$.

It remains to show that Y_{t+1,n_m} converges jointly. Let

$$g(\cdot) \in \mathrm{BL}_1 \equiv \left\{ \sup_x |g(x)| \le 1, |g(x) - g(y)| \le d(x,y) \right\}$$

be a bounded Lipschitz function with respect to the distance metric d, where d is the product metric for $\Pi_{1:t+1}, Y_{1:t+1}$:

$$d(G_{t+1}, G'_{t+1}) = \max_{s \in [t+1], k \in [K]} \left\{ |\Pi_{s,k} - \Pi'_{s,k}|, |Y_{s,k} - Y'_{s,k}| \right\}.$$

Let $U_{n_m} = \Pi_{t+1,n_m} Z_{n_m}$ for some $Z_{n_m} \sim \mathcal{N}(0, \Sigma(P_{n_m}))$ independently. Observe that

$$\begin{aligned}
&\left| \mathbb{E}[g(G_{t,n_m}, \Pi_{t+1,n_m}, Y_{t+1,n_m})] - \mathbb{E}[g(G_{t,n_m}, \Pi_{t+1,n_m}, U_{n_m})] \right| \\
&= \left| \mathbb{E}\left[\mathbb{E}[g(G_{t,n_m}, \Pi_{t+1,n_m}, Y_{t+1,n_m}) \mid G_{t,n_m}, \Pi_{t+1,n_m}] \right] - \mathbb{E}\left[\mathbb{E}[g(G_{t,n_m}, \Pi_{t+1,n_m}, U_{n_m}) \mid G_{t,n_m}, \Pi_{t+1,n_m}] \right] \right|
\end{aligned}$$

We claim that for some c_m independent of g,

$$\begin{aligned}
&\left| \mathbb{E}[g(G_{t,n_m}, \Pi_{t+1,n_m}, Y_{t+1,n_m}) \mid G_{t,n_m}, \Pi_{t+1,n_m}] - \mathbb{E}[g(G_{t,n_m}, \Pi_{t+1,n_m}, W_{n_m}) \mid G_{t,n_m}, \Pi_{t+1,n_m}] \right| \\
&\le c_m \to 0
\end{aligned} \tag{C.6}$$

with probability one. Assuming this claim, we would show that

$$\sup_{g \in \mathrm{BL}_1} \left| \mathbb{E}[g(G_{t,n_m}, \Pi_{t+1,n_m}, Y_{t+1,n_m})] - \mathbb{E}[g(G_{t,n_m}, \Pi_{t+1,n_m}, W_{n_m})] \right| \to 0.$$

Observe that by construction of U_{n_m} and (C.5), for some Y_{t+1},

$$(G_{t,n_m}, \Pi_{t+1,n_m}, U_{n_m}) \xrightarrow{d} (G_t, \Pi_{t+1}, Y_{t+1}).$$

Equivalently,

$$\sup_{g \in \mathrm{BL}_1} \left| \mathbb{E}[g(G_{t,n_m}, \Pi_{t+1,n_m}, U_{n_m})] - \mathbb{E}[g(G_t, \Pi_{t+1}, Y_{t+1})] \right| \to 0.$$

Therefore, by the triangle inequality,

$$\sup_{g \in \mathrm{BL}_1} \left| \mathbb{E}[g(G_{t,n_m}, \Pi_{t+1,n_m}, Y_{t+1,n_m})] - \mathbb{E}[g(G_t, \Pi_{t+1}, Y_{t+1})] \right| \to 0.$$

Equivalently,

$$G_{t+1,n_m} \xrightarrow{d} G_{t+1}.$$

Thus, to conclude the proof, we show (C.6). For given values of $G_{t,n_m}, \Pi_{t+1,n_m} = (G, \Pi)$, let

$$h(y) = g(G, \Pi, y).$$

Observe that $h(\cdot)$ is bounded and Lipschitz with respect to $d_y(y_1, y_2) = \max_k |y_{1k} - y_{2k}|$. Thus, (C.6) amounts to the following:

$$\left|\mathbb{E}_{Y_{t+1,n_m}|\Pi}[h(Y_{t+1,n_m})] - \mathbb{E}_{U_{n_m}|\Pi}[h(U_{n_m})]\right| \leq c_m.$$

Note that by Lemma C.2 and Assumption A.1, for prok the Prokhorov metric (see Lemma C.2 and Proposition A.5.2 in van der Vaart and Wellner (1996)),

$$\left|\mathbb{E}_{Y_{t+1,n_m}|\Pi}[h(Y_{t+1,n_m})] - \mathbb{E}_{U_{n_m}|\Pi}[h(U_{n_m})]\right| \lesssim \sup_{\Pi \in \Delta^{K-1}, P \in \mathcal{P}} \mathsf{prok}\left(Y_{n_m}(\Pi, P), U_{n_m}(\Pi, P)\right) \leq c_m \to 0,$$

where

$$Y_{k,n_m}(\Pi, P) = \frac{1}{\sqrt{n_t}} \sum_{i \in \mathcal{I}_t} D_{ik} X_i(k) \quad D_i \sim \text{Mult}(1, \Pi), X_i(\cdot) \sim P$$

and

$$U_{n_m}(\Pi, P) \sim \mathcal{N}(0, \Pi\Sigma(P)).$$

This concludes the proof. □

C.2.1. *Stopping time.* So far, we have shown that G_{T_0,n_m} converges. Now, we consider the following assumption on the stopping time T_n. Let $\Xi_{t,n} = \mathbb{1}(T_n > t)$. We assume that $\Xi_{t,n}$ satisfies conditions analogous to those satisfied by κ_t in Assumption C.1.

Assumption C.2. *We assume that for all $t = 1, \ldots, T_0$,*

(1) $\Xi_{t,n} = \Xi_{t,n}\left(X_{1:t-1,n}, \hat{\Pi}_{1:t-1,n}, \hat{\Sigma}_{t-1,n}, n_{1:t-1}\right) = \Xi_{t,n}\left(Y_{1:t-1,n}, \hat{\Pi}_{1:t-1,n}, \hat{\Sigma}_{t-1,n}, n_{1:t-1}, \mu_n\right)$ *is measurable with respect to* $(Y_{1:t-1,n}, \hat{\Pi}_{1:t-1,n})$ *and* $\Xi_{T_0,n} = 0$ *a.s.*

(2) Let P_{n_m} denote a subsequence of the data-generating process satisfying the conclusion of Lemma C.1. Then, if

$$(\hat{\Pi}_{1:T_0,n}, Y_{1:T_0,n}) \xrightarrow{d} G_{T_0} = (\Pi_{1:T_0}, Y_{1:T_0})$$

then for all $t = 1, \ldots, T_0$, we have the following convergences:

$$\Xi_{t,n_m}\left(Y_{1:t-1,n}, \hat{\Pi}_{1:t-1,n}, \hat{\Sigma}_{t-1,n}, n_{1:t-1}, \mu_n\right) \xrightarrow{d} \Xi_t\left(Y_{1:t-1}, \Pi_{1:t-1}; h, \sigma^2_{1:K}, c_{1:t-1}\right)$$

and

$$T_n \xrightarrow{d} T \equiv \underset{t=1,\ldots,T_0}{\arg\max}\left(\Xi_{t-1} - \Xi_t\right)$$

jointly.

(3) For all $t = 1, \ldots, T_0 - 1$, the limiting random variable $\Xi_t\left(Y_{1:t-1}, \Pi_{1:t-1}; h, \sigma^2_{1:K}, c_{1:t-1}\right)$ is measurable with respect to the same random variables as Π_t is in Assumption C.1(3). That is, Ξ_t is measurable with respect to $\{B_s \mathbf{Y}_{1:s} : s \leq t-1\}$ for matrices B_s in Assumption C.1(3).

Note that we can write

$$L_n - \lambda_n' \mu = \sum_{s=1}^{T_0} \underbrace{(\Xi_{s-1,n} - \Xi_{s,n})}_{\mathbb{1}(T_n = s)} \sum_{t=1}^{s-1} \sum_{k=1}^{K} \frac{n_t \hat{\Pi}_{tkn}}{\sqrt{n} \hat{\sigma}_{skn}^2} (X_{tkn} - \mu)$$

and

$$Y_{T_n,n} = \sum_{s=1}^{T_0} (\Xi_{s-1,n} - \Xi_{s,n}) Y_{sn}.$$

The following corollary shows the asymptotic analogue of Theorem 3.7.

Corollary C.1. *Additionally under Assumption C.1(3) and Assumption C.2, on the subsequence n_m,*

$$\begin{bmatrix} L_{n_m} - \lambda_n' \mu_{n_m} \\ Y_{T,n_m} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} 1'\Sigma^{-1} Y_{:T-1} \\ Y_T \end{bmatrix},$$

where, in the limit,

$$\begin{bmatrix} 1'\Sigma^{-1} Y_{:T-1} \\ Y_T \end{bmatrix} \mid (T, \{B_s \mathbf{Y}_{1:s} : s \le T-1\}) \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1'\Sigma^{-1}\Pi_{:T-1} & 0 \\ 0 & \Sigma \Pi_T \end{bmatrix} \right)$$

Proof. The weak convergence is by the continuous mapping theorem and Lemma C.4, noting that $\hat{\sigma}_{skn} \to \sigma_k^2$ and $n_t/n \to c_t$ for all s, t.

The second claim is a claim on the properties of the limiting distribution. Note that Assumption C.1(3) and Assumption C.2(3) ensure that the limiting assignment probabilities and the limiting stopping time are functions of certain differences in $Y_{1:t}$. As a result, we can apply Lemma B.1 directly to finish the proof. □

C.3. **Behavior of the test statistic.** We now rewrite the empirical analogue of the least-squares coefficient (A.2):

$$S_n^\star = \left(\frac{\lambda_n \lambda_n' / \sqrt{n}}{\lambda_n' 1} + \frac{n_{T_n}}{n} \hat{\Sigma}_n^{-1} \hat{\Pi}_{T_n n} \right)^+ \left(\frac{\lambda_n}{\lambda_n' 1} \frac{L_n}{\sqrt{n}} + \frac{n_{T_n}}{n} \hat{\Sigma}_n^{-1} \hat{\Pi}_{T_n n} X_{T_n n} \right).$$

Let $\eta_n = \eta(\hat{\Pi}_{1:T_n,n})$ be a direction of inference. We additionally assume that when the experiment stops, the probability of the last batch is pruned so as to exclude $(0, \epsilon)$.

Assumption C.3 (Pruning)**.** *For some $\epsilon > 0$, the random variable*

$$\Pi_{T_n,n} = \kappa_{T_n,n}(\cdot) = \sum_{s=1}^{T} (\Xi_{s-1,n} - \Xi_{s,n}) \kappa_{s,n}(\cdot)$$

takes values outside of $\{q \in \Delta^{K-1} : q_k \in (0, \epsilon) \text{ for some } k \in [K]\}$ with probability one.

We assume that $\eta(\cdot)$ is suitably continuous.

Assumption C.4. *Under a subsequence n_m of n that satisfies the consequences of Lemma C.1 and Theorem C.1, if*

$$(\hat{\Pi}_{1:T_0,n}, Y_{1:T_0,n}) \xrightarrow{d} G_{T_0} \text{ and } T_n \xrightarrow{d} T$$

Then $\eta_n = \eta\left(Y_{1:T_n,n}, \hat{\Pi}_{1:T_n,n}, T_n, \hat{\Sigma}_{T_n,n}, n_{1:T_n}, \mu_n\right) \xrightarrow{d} \eta(Y_{1:T}, \Pi_{1:T}, T, \Sigma, c_{1:T})$ *along the subsequence n_m. Moreover, η is measurable with respect to $(T, \{B_s \mathbf{Y}_{1:s} : s \le T-1\})$ for matrices B in Assumption A.3. We also assume that $\|\eta_{n_m}\| > \epsilon > 0$ with probability one. Moreover, $\eta(T, \Pi_{1:T})$ is in the column space of*

$$\frac{\lambda\lambda'}{1'\lambda} + c_T \Sigma^{-1} \Pi_T$$

almost surely and η_{n_m} is in the column space of

$$\frac{\lambda_{n_m}\lambda_{n_m}'/\sqrt{n_m}}{\lambda_{n_m}' 1} + \frac{n_T}{n_m} \hat{\Sigma}_{n_m}^{-1} \hat{\Pi}_{T n_m}$$

almost surely for all sufficiently large m.

Under this assumption, note that

$$\sqrt{n}(\eta_n' S_n^\star - \eta_n' \mu_n) = \eta_n' \left(\frac{\lambda_n \lambda_n'/\sqrt{n}}{\lambda_n' 1} + \frac{n_T}{n} \hat{\Sigma}_n^{-1} \hat{\Pi}_{Tn}\right)^+ \left(\frac{\lambda_n}{\lambda_n' 1}(L_n - \lambda_n' \mu_n) + \sqrt{c_T} \hat{\Sigma}_n^{-1} Y_{Tn}\right)$$

since $\eta_n' Q^+ Q \mu = \eta_n' \mu$ if η_n is in the column space of the symmetric matrix Q. Let

$$\hat{\sigma}_{\tau,n}^2 = \eta_n' \left(\frac{\lambda_n \lambda_n'/\sqrt{n}}{\lambda_n' 1} + \frac{n_T}{n} \hat{\Sigma}_n^{-1} \hat{\Pi}_{Tn}\right)^+ \eta_n.$$

Then, we obtain the uniform convergence of the test statistic.

Theorem C.2. *Under Assumptions A.1 and C.1 to C.4, for every sequence $P_n \in \mathcal{P}$ and every subsequence n_r of n, there exists a further subsequence n_m such that, jointly,*

$$\frac{\sqrt{n_m}(\eta_{n_m}' S_{n_m}^\star - \eta_{n_m}' \mu_{n_m})}{\hat{\sigma}_{\tau,n_m}} \xrightarrow{d} Z \quad (\eta_{n_m}, \Pi_{1:T_{n_m}}, T_{n_m}) \xrightarrow{d} (\eta, \Pi_{1:T}, T)$$

where $Z \mid (\eta, \Pi_{1:T}, T) \sim \mathcal{N}(0,1)$.

Proof. Let n_m be the subsequence that satisfies Corollary C.1. Under Assumption C.4, $\eta_{n_m} \xrightarrow{d} \eta$ jointly with $(\Pi_{1:T_{n_m}}, T_{n_m}) \xrightarrow{d} (\Pi_{1:T}, T)$. By the continuous mapping theorem and Lemma C.4,

$$\frac{\lambda_{n_m}\lambda_{n_m}'/\sqrt{n_m}}{\lambda_{n_m}' 1} + \frac{n_{T_{n_m}}}{n_m} \hat{\Sigma}_{n_m}^{-1} \hat{\Pi}_{T_{n_m} n_m} \xrightarrow{d} \frac{\lambda'\lambda}{\lambda' 1} + c_T \Sigma^{-1} \Pi_T$$

jointly with $G_{T_{n_m}, n_m} \xrightarrow{d} G_{T, n_m}$.

Next, Lemma C.6 shows that the convergence in the above display is preserved by the Moore-Penrose pseudoinverse, despite the discontinuity in the pseudoinverse. This is becasue of pruning.

Next, we show that the standard error is positive in the limit. Note that there exists some $c > 0$ such that, uniformly over $\mathcal{P}$,

$$\eta' \left(\frac{\lambda'\lambda}{\lambda' 1} + c_T \Sigma^{-1} \Pi_T \right)^{+} \eta \geq \lambda_{\max}^{-1} \left(\frac{\lambda'\lambda}{\lambda' 1} + c_T \Sigma^{-1} \Pi_T \right) \|\eta\| > c > 0$$

where $\lambda_{\max}(A)$ is the maximum eigenvalue of A. Thus, continuous mapping theorem implies that

$$\frac{\sqrt{n}(\eta_n' S_n^\star - \eta_n' \mu_n)}{\hat{\sigma}_{\tau,n}} \xrightarrow{d} \frac{\eta' \left(\frac{\lambda'\lambda}{\lambda' 1} + c_T \Sigma^{-1} \Pi_T \right)^{+} \left(\frac{\lambda}{\lambda' 1} 1' \Sigma^{-1} Y_{:T-1} + \sqrt{c_T} \Sigma^{-1} Y_T \right)}{\sqrt{\eta' \left(\frac{\lambda'\lambda}{\lambda' 1} + c_T \Sigma^{-1} \Pi_T \right)^{+} \eta}} \equiv Z.$$

Note that by Corollary C.1

$$\eta' \left(\frac{\lambda'\lambda}{\lambda' 1} + c_T \Sigma^{-1} \Pi_T \right)^{+} \left(\frac{\lambda}{\lambda' 1} 1' \Sigma^{-1} Y_{:T-1} + \sqrt{c_T} \Sigma^{-1} Y_T \right) \mid \eta, \Pi_{1:T}, T$$
$$\sim \mathcal{N}\left(0, \eta' \left(\frac{\lambda'\lambda}{\lambda' 1} + c_T \Sigma^{-1} \Pi_T \right)^{+} \eta \right)$$

Thus,

$$\frac{\eta' \left(\frac{\lambda'\lambda}{\lambda' 1} + c_T \Sigma^{-1} \Pi_T \right)^{+} \left(\frac{\lambda}{\lambda' 1} 1' \Sigma^{-1} Y_{:T-1} + \sqrt{c_T} \Sigma^{-1} Y_T \right)}{\sqrt{\eta' \left(\frac{\lambda'\lambda}{\lambda' 1} + c_T \Sigma^{-1} \Pi_T \right)^{+} \eta}} \mid \eta, \Pi_{1:T}, T \sim \mathcal{N}(0,1).$$

This completes the proof since $\eta, \Pi_{1:T}, T$ is measurable with respect to

$$(T, \{B_s \mathbf{Y}_{1:s} : s \leq T-1\}).$$

□

C.4. **Conditional coverage performance.**

Assumption C.5. *Let $f(\Pi_{1:T}, T, \eta) < C < \infty$ be a bounded function such that along all subsequences $P_{n_m} \in \mathcal{P}$, if*

$$(\hat{\Pi}_{1:T_{n_m}, n_m}, T_{n_m}, \eta_{n_m}) \xrightarrow{d} (\Pi_{1:T}, T, \eta)$$

then

$$f(\hat{\Pi}_{1:T_{n_m}, n_m}, T_{n_m}, \eta_{n_m}) \xrightarrow{d} f\left(\Pi_{1:T}, T, \eta\right).$$

Theorem C.3. *Suppose Assumptions A.1 and C.1 to C.4 hold and let $\alpha \in (0, 0.5)$. Let $f(\Pi_{1:T}, T, \eta) > 0$ be a function satisfying Assumption C.5. Then,*

$$\limsup_{n\to\infty} \sup_{P\in\mathcal{P}} \left| \mathbb{E}_P \left[\mathbb{1}\left(|\eta_n' S_n^\star - \eta_n' \mu(P)| > z_{1-\alpha/2} \frac{\hat{\sigma}_{\tau,n}}{\sqrt{n}} \right) f(\Pi_{1:T_n,n}, T_n, \eta_n) \right] - \alpha \mathbb{E}_P[f(\Pi_{1:T_n,n}, T_n, \eta_n)] \right| = 0.$$

where $z_{1-\alpha/2} = \Phi^{-1}(1-\alpha/2)$.

Proof. Suppose to the contrary, for some $\epsilon > 0$,

$$\limsup_{n\to\infty} \sup_{P\in\mathcal{P}} \left| \mathbb{E}_P \left[\mathbb{1}\left(|\eta_n' S_n^\star - \eta_n'\mu(P)| > z_{1-\alpha/2}\frac{\hat{\sigma}_{\tau,n}}{\sqrt{n}} \right) f(\Pi_{1:T_n,n}, T_n, \eta_n) \right] - \alpha\mathbb{E}_P[f(\Pi_{1:T_n,n}, T_n, \eta_n)] \right| > 2\epsilon.$$

Then there exists a subsequence P_{n_r} where for all $r > 0$,

$$\limsup_{n\to\infty} \sup_{P\in\mathcal{P}} \left| \mathbb{E}_P \left[\mathbb{1}\left(|\eta_n' S_n^\star - \eta_n'\mu(P)| > z_{1-\alpha/2}\frac{\hat{\sigma}_{\tau,n}}{\sqrt{n}} \right) f(\Pi_{1:T_n,n}, T_n, \eta_n) \right] - \alpha\mathbb{E}_P[f(\Pi_{1:T_n,n}, T_n, \eta_n)] \right| > \epsilon. \tag{C.7}$$

By Theorem C.2, we can extract a further subsequence n_m along which

$$\sqrt{n_m}\frac{|\eta_{n_m}' S_{n_m}^\star - \eta_{n_m}'\mu(P_{n_m})|}{\hat{\sigma}_{\tau,n_m}} \xrightarrow{d} Z \tag{C.8}$$

and

$$(\eta_{n_m}, \Pi_{1:T_{n_m}}, T_{n_m}) \xrightarrow{d} (\eta, \Pi_{1:T}, T)$$

where

$$Z \mid (\eta, \Pi_{1:T}, T) \sim \mathcal{N}(0,1).$$

By Assumption C.5, we additionally have

$$f(\Pi_{1:T_{n_m},n_m}, T_{n_m}, \eta_{n_m}) \xrightarrow{d} f(\Pi_{1:T}, T, \eta).$$

Thus

$$\mathbb{1}\left(\sqrt{n_m}\frac{|\eta_{n_m}' S_{n_m}^\star - \eta_{n_m}'\mu(P_{n_m})|}{\hat{\sigma}_{\tau,n_m}} < z_{1-\alpha/2} \right) f(\Pi_{1:T_{n_m},n_m}, T_{n_m}, \eta_{n_m}) \xrightarrow{d} \mathbb{1}(Z < z_{1-\alpha/2}) f(\Pi_{1:T}, T, \eta).$$

Since f is bounded, we have that the expectations also converge. Note that

$$\mathbb{E}[\mathbb{1}(Z < z_{1-\alpha/2}) f(\Pi_{1:T}, T, \eta)] = \alpha\mathbb{E}[f(\Pi_{1:T}, T, \eta)].$$

Therefore, along this subsequence n_m, the limit is zero, contradicting (C.7). □

C.5. **Auxiliary lemmas.**

Lemma C.2 (Uniform central limit theorem via Lyapunov condition)**.** *Let $\pi \in \Delta^{K-1}$ be a vector of probabilities. Let $X_i, \ldots, X_n \overset{\text{i.i.d.}}{\sim} Q$ be $\mathbb{R}^K$-valued random vectors for some distribution $Q \in \mathcal{Q}$ in a family $\mathcal{Q}$, with $\mathbb{E}_Q[X] = \mu(Q)$. Let $\Sigma(Q) = \operatorname{diag}(\operatorname{Var}_Q(X_1), \ldots, \operatorname{Var}_Q(X_K))$. Let $Y(Q,\pi) \in \mathbb{R}^K$ be such that its k^{th} coordinate is*

$$Y_k(Q,\pi) = \frac{1}{\sqrt{n}}\sum_{i=1}^n D_{ik}(X_{ik} - \mu_k(Q)) \quad D_i = [D_{i1}, \ldots, D_{iK}]' \overset{\text{i.i.d.}}{\sim} \operatorname{Mult}(1,\pi).$$

Correspondingly, write $Y(Q,\pi) = \frac{1}{\sqrt{n}}\sum_{i=1}^n Y_i$. Suppose that there exists $c > 0$ such that

$$\sup_{Q\in\mathcal{Q}} \mathbb{E}_Q\left[\|X\|^{2+c}\right] < C < \infty,$$

where $\|\cdot\|$ *denotes the Euclidean norm. Then, there exists* $c_n \to 0$ *dominating the* Prokhorov metric*:*

$$\sup_{Q \in \mathcal{Q}} \sup_{\pi \in \Delta^{K-1}} \mathsf{prok}(Y(Q,\pi), \mathcal{N}(0, \Sigma(Q)\Pi)) \le c_n \to 0,$$

where prok *denotes the Prokhorov metric and* $\Pi = \operatorname{diag}(\pi)$*. Since the Prokhorov metric dominates the bounded Lipschitz metric, the same statement hold for the latter.*

Proof. This is a straightforward application of Proposition A.5.2 in van der Vaart and Wellner (1996). Observe that, since the entries of D_i are bounded by 1, Y_i satisfies the Lyapunov condition

$$\|Y_i\| \le \|X_i\| \text{ almost surely } \implies \mathbb{E}_{Q,\pi}\|Y_i\|^{2+c} < C$$

and $\operatorname{Var}_{Q,\pi}(Y_i) = \Sigma(Q) \operatorname{diag}(\pi)$.

For a given Q, let

$$\begin{aligned}
g(t, Q, \pi) &= \mathbb{E}_{Q,\pi}\|Y_i\|^2 \mathbb{1}(\|Y_i\| \ge t) \\
&\le \mathrm{P}_{Q,\pi}(\|Y_i\| \ge t) \mathbb{E}_{Q,\pi}[\|Y_i\|^{2+c}] \\
&\le C \cdot \mathrm{P}_{Q,\pi}(\|Y_i\| \ge t) \\
&\le C \cdot \left(\frac{\mathbb{E}[\|Y_i\|^{2+c}]}{t^{2+c}} \wedge 1 \right) = \frac{C^2}{t^{2+c}} \wedge C
\end{aligned}$$

where the first inequality is via Hölder's inequality, the second inequality is via the Lyapunov condition, and the third inequality is via Markov's inequality combined with the Lyapunov condition.

Then Proposition A.5.2 in van der Vaart and Wellner (1996) implies that there are constants $C_1, \ldots, C_5$, dependent only on the moment bound C, such that for any δ,

$$\begin{aligned}
\mathsf{prok}\left(Y(Q,\pi), \mathcal{N}(0, \Sigma(Q,\pi))\right) &\le \delta + C_1 g(\delta\sqrt{n}, Q, \pi)^{1/3} + C_2 \delta^{1/4} \left(1 + C_3 + C_4 \sqrt{|\log \delta|}\right) \\
&\le C_5 \left(\delta + \left(1 \wedge \frac{1}{\delta^{2+c} n^{1+c/2}}\right)^{1/3} + \delta^{1/4} \sqrt{|\log \delta|} \right)
\end{aligned}$$

We can pick $c_n \to 0$ to be the expression on the RHS with, say, $\delta = n^{-1/4}$ to complete the proof. ∎

Lemma C.3. *Let* $N \sim \operatorname{Bin}(n, p)$*, then with probability at least* $1 - \Delta_n$*,*

$$\left| \frac{N}{n} - p \right| \le d_n$$

where $d_n, \Delta_n \to 0$ *are sequences independent of* p*. Hence, by a union bound, for any* K *variables* $N_1, \ldots, N_K$ *where* $N_k \sim \operatorname{Bin}(n, p_k)$*,*

$$\mathrm{P}\left[\forall k : \left| \frac{N_k}{n} - p_k \right| \le d_n \right] \ge 1 - K\Delta_n.$$

Proof. Exercise 2.3.5 in Vershynin (2018) shows the following Chernoff bound: For some absolute $c > 0$,

$$\mathrm{P}\left[\left|\frac{N}{n} - p\right| \geq \delta p\right] \leq 2e^{-cnp\delta^2}$$

If we choose $\delta = d_n/p$, we have

$$\mathrm{P}\left[\left|\frac{N}{n} - p\right| \geq d_n\right] \leq 2e^{-cnd_n^2/p} \leq 2e^{-cnd_n^2}$$

Pick, for instance, $d_n = \log(n)/\sqrt{n} + 1/n$ and $\Delta_n = 1 - 2n^{-c}$ to complete the proof. $\square$

Lemma C.4. *Under our setup, consider a subsequence n_m. Let V_{tn_m} be certain random variables indexed by $t = 1, \ldots, T_0$ and n_m. Suppose*

$$(V_{1n_m}, \ldots, V_{T_0,n_m}, \Xi_{1n_m}, \ldots, \Xi_{T_0,n_m}, T_{n_m}) \xrightarrow{d} (V_1, \ldots, V_{T_0}, \Xi_1, \ldots, \Xi_{T_0}, T)$$

Then

$$V_{T_{n_m},n_m} \xrightarrow{d} V_T$$

Proof. Note that, by continuous mapping theorem,

$$V_{T_{n_m},n_m} = \sum_{t=1}^{T_0} (\Xi_{t-1,n_m} - \Xi_{t,n_m}) V_{t,n_m} \xrightarrow{d} \sum_{t=1}^{T_0} (\Xi_{t-1} - \Xi_t) V_t = V_T.$$

This concludes the proof. $\square$

Lemma C.5. *Let M_n, M be random positive semi-definite matrices such that $M_n \xrightarrow{d} M$. Suppose that there exists sets $\mathcal{M}_n, \mathcal{M}$ such that (a) for all n, $\mathrm{P}(M_n \in \mathcal{M}_n) = \mathrm{P}(M \in \mathcal{M}) = 1$ and that (b) for every deterministic sequence $v_n \to v$ where $v_n \in \mathcal{M}_n$ and $v \in \mathcal{M}$, we have that for all sufficiently large n $\mathrm{rank}(v_n) = \mathrm{rank}(v)$. Then $M_n^+ \xrightarrow{d} M^+$ where $(\cdot)^+$ is the Moore-Penrose pseudo-inverse.*

Proof. Since $\mathrm{rank}(v_n) = \mathrm{rank}(v)$ for every sequence eventually, we have that $v_n^+ \to v^+$ for all sequences (Rakočević, 1997). Hence by the extended continuous mapping theorem, $M_n^+ \xrightarrow{d} M^+$. $\square$

Lemma C.6. *Under the conditions of Theorem C.2 and in its proof,*

$$\left(\frac{\lambda_{n_m}\lambda_{n_m}'/\sqrt{n_m}}{\lambda_{n_m}'1} + \frac{n_{T_{n_m}}}{n_m}\hat{\Sigma}_{n_m}^{-1}\hat{\Pi}_{T_{n_m}n_m}\right)^+ \xrightarrow{d} \left(\frac{\lambda'\lambda}{\lambda'1} + c_T\Sigma^{-1}\Pi_T\right)^+$$

jointly.

Proof. We verify this statement with Lemma C.5. Let

$$M_{n_m} = \frac{\lambda_{n_m}\lambda_{n_m}'/\sqrt{n_m}}{\lambda_{n_m}'1} + \frac{n_{T_{n_m}}}{n_m}\hat{\Sigma}_{n_m}^{-1}\hat{\Pi}_{T_{n_m}n_m}$$

and $M = \frac{\lambda'\lambda}{\lambda'1} + c_T \Sigma^{-1}\Pi_T$ where $M_n \xrightarrow{d} M$ (along a subsequence) is already shown. Consider a fixed sequence $v_m \to v$. Note that v_m is of the form

$$\frac{\ell_m \ell_m' / \sqrt{n_m}}{\ell_m' 1} + \frac{n_{T_m}}{n_m} \Omega_m^{-1} \pi_{T_m, m}$$

Similarly, we can write

$$v = \frac{\ell\ell'}{\ell' 1} + c_{T^*} \Omega^{-1} \pi_{T^*}$$

Note that $v_m \to v$ implies that $\frac{\ell_m \ell_m'/\sqrt{n_m}}{\ell_m' 1} \to \frac{\ell\ell'}{\ell'1}$, by inspecting the convergence of the off-diagonal elements. Hence

$$\frac{n_{T_m}}{n_m} \Omega_m^{-1} \pi_{T_m, m} \to c_{T^*} \Omega^{-1} \pi_{T^*}$$

Since $\Pi_{T_{n_m}, n_m} \xrightarrow{d} \Pi_T$, by Assumption C.3, we can restrict to considering π_{T^*} whose entries are never within $(0, \epsilon)$. Thus the diagonal entries of $c_{T^*}\Omega^{-1}\pi_{T^*}$ are never within $(0, \min_{t \in [T_0]} c_t \frac{1}{C_1} \epsilon)$ by Assumption A.1. Hence, if

$$\frac{\ell_m \ell_m' / \sqrt{n_m}}{\ell_m' 1} \to \frac{\ell\ell'}{\ell' 1} \quad \frac{n_{T_m}}{n_m} \Omega_m^{-1} \pi_{T_m, m} \to c_{T^*} \Omega^{-1} \pi_{T^*},$$

then for all sufficiently large m, we must have that (a) the number and location of zeros in the diagonal of $\frac{n_{T_m}}{n_m} \Omega_m^{-1} \pi_{T_m, m}$ remain fixed and (b) ℓ_m has strictly positive entries.

It suffices to verify that for a strictly positive vector $a \in \mathbb{R}^K$ and diagonal matrix $A \succeq 0$ with rank r, $aa' + A$ has rank equal to $K \wedge (r+1)$. Given this claim, since the number of zeros in the diagonal of $\frac{n_{T_m}}{n_m} \Omega_m^{-1} \pi_{T_m, m}$ remain unchanged for sufficiently large m, its rank is constant. Thus the rank of v_m is also constant for sufficiently large m.

To verify this claim, we will show that a and $\{e_k : A_{kk} \neq 0\}$ span the column space of $aa' + A$, where e_k is the k^{th} standard basis vector. First, note that if $A \succ 0$ is positive definite, then $aa' + A$ is full-rank, and the claim is true. Otherwise, there is some k for which $A_{kk} = 0$. Then $(aa' + A)e_k = (a'e_k)a$. Since a has strictly positive entries, this shows that a is in the column space of $aa' + A$. Let ℓ be such that $A_{\ell\ell} \neq 0$. Then $(aa' + A)e_\ell = A_{\ell\ell} e_\ell + (a'e_\ell)a$ is a linear combination of e_ℓ and a. Since a is in the column space of $aa' + A$, then e_ℓ is also in the column space. Lastly, since every vector in the column space is in the span of a and $\{e_k : A_{kk} \neq 0\}$: $(aa' + A)u = (a'u)a + \sum_{k: A_{kk} \neq 0} A_{kk} u_k e_k$, we conclude that a and $\{e_k : A_{kk} \neq 0\}$ span the column space of $aa' + A$. We finish the proof by noting that the dimension of a and $\{e_k : A_{kk} \neq 0\}$ is $K \wedge (r+1)$. $\square$

C.6. **Verification of Theorem A.2.**

Theorem A.2. *Under Assumptions A.1 to A.5, level-$(1-\alpha)$ two-sided confidence intervals $\mathrm{CS}_n(\alpha) \equiv \eta_n' S_n^\star \pm \Phi^{-1}(1-\alpha/2) \cdot \frac{\hat\sigma_{\tau,n}^2}{\sqrt{n}}$ have exact conditional asymptotic size for the parameters*

τ_n: *For all* η, T,

$$\liminf_{n\to\infty} \inf_{P\in\mathcal{P}} \left| P\left(\tau_n \in \mathrm{CS}_n(\alpha) \mid \eta_n = \eta, T_n = T\right) - \alpha \right| P(\eta_n = \eta, T_n = T) = 0.$$

Proof. The theorem follows from Theorem C.3 pending verification of Assumptions C.1 to C.5.

Note that our setup means that Assumption C.1(1) and Assumption C.2(1) are satisfied. Lemma C.7 verifies Assumption C.1(2–3) and Assumption C.2(2–3). Assumption A.4 implies Assumption C.3 directly. Lemma C.8 verifies Assumption C.4. Finally, since η_n and T_n both have finite support (with discrete topology), for any η_1, T_1 in the support, $f(\eta, T) = \mathbb{1}(\eta = \eta_1, T = T_1)$ satisfies Assumption C.5. □

Lemma C.7. *If the assignment algorithm satisfies Assumptions A.2 and A.3 then it satisfies Assumption C.1(2–3) and Assumption C.2(2–3).*

Proof. Note that we can rewrite

$$\sqrt{n} W_{t,n} = \underbrace{\left(\sum_{s=1}^{t} \frac{n_t}{n} \hat{\Pi}_{sn}\right)^{-1}}_{\hat{\Pi}_{:t,n}^{-1}} \underbrace{\sum_{s=1}^{t} \sqrt{\frac{n_t}{n}} Y_{sn}}_{Y_{:t,n}} + \underbrace{\sqrt{n}\mu_n}_{h_n + \max_\ell \mu_{n,\ell}}$$

Hence we can equivalently write

$$\kappa_{t+1}\left(\sqrt{n} W_{t,n}, \hat{\Sigma}_{tn} \hat{\Pi}_{:t,n}^{-1}\right) = \kappa_{t+1}\left(\hat{\Pi}_{:t,n}^{-1} Y_{:t,n} + h_n, \hat{\Sigma}_{tn} \hat{\Pi}_{:t,n}^{-1}\right)$$

and similarly for Ξ_{t+1}.

With respect to the subsequence P_{n_m}, let $J = \{j : h_{n_m,j} > -\infty\}$. We restrict to the subsequence n_m and index by n instead.

To verify Assumption C.1(2–3) and Assumption C.2(2–3), we induct on t. We will verify that

- The statements in Assumption C.1(2–3) and Assumption C.2(2–3) hold for batch t
- $\mathrm{P}(\Pi_{t+1,j} > 0 \mid \Xi_t = 1) = 1$ for all $j \in J$.

The base case is $t = 1$ for κ_2, Ξ_2. Note that $\hat{\Sigma}_{tn} \xrightarrow{p} \Sigma > 0$. Under (C.2), $\hat{\Pi}_{:1,n} = \frac{n_1}{n} \hat{\Pi}_{1,n} \xrightarrow{p} c_1 \Pi_1 > 0$. Thus,

$$\begin{bmatrix} Y_{:1,n} \\ \hat{\Pi}_{:1,n} \\ \hat{\Sigma}_{tn} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} \sqrt{c_1} Y_1 \\ c_1 \Pi_1 \\ \Sigma \end{bmatrix}.$$

Hence, for $t = 1$,

$$\begin{bmatrix} \hat{\Pi}_{:t,n}^{-1} Y_{:t,n} + h_n \\ \hat{\Sigma}_{tn} \hat{\Pi}_{:t,n}^{-1} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} (c_1)^{-1/2} \Pi_1^{-1} Y_1 + h \\ \Sigma \Pi_1^{-1} \end{bmatrix}.$$

Here, the limit $(c_1)^{-1/2}\Pi_1^{-1}Y_1 + h$ is absolutely continuous with respect to the Lebesgue measure on indices in J, since Y_1 is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^K$ and $\Pi_1 > 0$.

Hence Assumptions A.2 and A.3 imply that the set of discontinuities of $\Xi_2(\cdot), \kappa_{2,(1)}(\cdot,\cdot)$ and $\kappa_{2,(0)}$ is measure zero with respect to the law of $((c_1)^{-1/2}\Pi_1^{-1}Y_1 + h, \Sigma\Pi_1^{-1})$. Thus, the continuous mapping theorem applies and κ_2, Ξ_2 converges. This verifies the statements in Assumption C.1(2) and Assumption C.2(2) for $t = 1$.

By location-invariance, Π_2, Ξ_2 are measurable with respect to the differences

$$\left\{ c_1^{-1/2}\Pi_{1k}Y_{1k} - c_1^{-1/2}\Pi_{1K}Y_{1K} : k \in [K-1] \right\}.$$

These differences can be written as $B_1 Y_1$ where $B_1\sqrt{c_1}\Pi_1 1_K = 0$. This proves Assumption C.1(2–3) and Assumption C.2(2–3) for $t = 1$.[7] Moreover, observe that $\mathrm{P}(\Pi_{2k} > 0) = 1$ whenever $h_{nk} > -\infty$.

In the inductive case, for a generic t, assume (C.3). We have that

$$\begin{bmatrix} \hat{\Pi}_{:t,n}^{-1} Y_{:t,n} + h_n \\ \hat{\Sigma}_{tn}\hat{\Pi}_{:t,n}^{-1} \end{bmatrix} \xrightarrow{d} \begin{bmatrix} \Pi_{:t}^{-1} Y_{:t} + h \\ \Sigma\Pi_{:t}^{-1} \end{bmatrix}.$$

Note that $\Pi_{:t} > c_1\Pi_1 > 0$ almost surely. Note that the inductive hypothesis implies that $\mathrm{P}(\Pi_{tk} > 0 \mid \Xi_{t-1} = 1) = 1$ for k where $h_{nk} > -\infty$. Note that this means $\{Y_{:t,j} : j \in J\}$ is absolutely continuous *conditional on* $\Pi_{1:t}, \Xi_t = 1$ with respect to the Lebesgue measure on $\mathbb{R}^{|J|}$, since the last term $Y_t \sim \mathcal{N}(0, \Pi_t\Sigma)$. As a result, the set of discontinuities of $\Xi_{t+1}, \kappa_{t+1,(0)}, \kappa_{t+1,(1)}$ is measure zero with respect to the law of $(\Pi_{:t}^{-1}Y_{:t} + h, \Sigma\Pi_{:t}^{-1}) \mid \Xi_t = 1$. Note that we can write

$$\Xi_{t+1} = \Xi_t \Xi_{t+1}.$$

Since both Ξ_{t+1} and κ_{t+1} are continuous when $\Xi_t = 0$, we conclude that their discontinuities are measure zero with respect to the law of $(\Pi_{:t}^{-1}Y_{:t} + h, \Sigma\Pi_{:t}^{-1}, \Xi_t)$. Therefore, κ_{t+1}, Ξ_{t+1} converges weakly by the continuous mapping theorem.

Similarly, by location-invariance, Π_{t+1} is measurable with respect to

$$\left\{ \Pi_{:t,k}^{-1} Y_{:t,k} - \Pi_{:t,K}^{-1} Y_{:t,K} : k \in [K-1] \right\}.$$

[7] It is easy to show that

$$T_n = \arg\max_t \Xi_{t-1,n} - \Xi_{t,n} \xrightarrow{d} T = \arg\max_t \Xi_{t-1} - \Xi_t$$

provided that $\Xi_{1:T_0,n} \xrightarrow{d} \Xi_{1:T_0}$ by the continuous mapping theorem, since the support of $(\Xi_{t-1,n} - \Xi_{t,n})_{t=2}^{T_0}$ is the set of standard basis vectors in $\mathbb{R}^{T_0}$.

When written as a transformation $B_t \mathbf{Y}_{1:t}$, we have that

$$B_t \begin{bmatrix} \sqrt{c_1}\Pi_1 1_K \\ \vdots \\ \sqrt{c_t}\Pi_t 1_K \end{bmatrix} = 0.$$

Lastly, we have that $\mathrm{P}(\Pi_{t+1,j} > 0 \mid \Xi_t = 1) = 1$ by assumption. This completes the proof. $\square$

Lemma C.8. *Under Assumption C.3 and Assumption A.1, if $\eta_n(\cdot)$ satisfies Assumption A.5, then it satisfies Assumption C.4.*

Proof. Let $\varsigma_n = \varsigma(W_{T_n,n})$ denote the strict ranking of the entries of $W_{T_n,n}$. Likewise, let $E_n = \{k : \Pi_{T_n,n,k} = 0\}$.

For the weak convergence of η_n along a subsequence n_m, it suffices to show that

$$(T_{n_m}, \varsigma_{n_m}, E_{n_m})$$

converges weakly along the subsequence, since their support is finite (and hence any function is continuous under the discrete topology). Note that since ς_n is a function of the indicators $I_{k,k',n} = \mathbb{1}(\sqrt{n}W_{T_n,t,k} - \sqrt{n}W_{T_n,t,k'} > 0), k \neq k'$ and $k, k' \in [K]$ (and a continuous function since the domain is discrete). It further suffices to show joint convergence of

$$(T_{n_m}, \{I_{k,k',n_m} : k, k' \in [K], k \neq k'\}, E_{n_m}).$$

Let us consider the above display as some function ψ_m of $U_m = (T_{n_m}, \hat{\Pi}_{T_{n_m}}, \sqrt{n_m}(W_{T_{n_m},n_m} - \mu_{n_m}))$, where

$$U_m \xrightarrow{d} (T, \Pi_T, W_T) \quad W_T = \left(\sum_{t=1}^{T} c_t \Pi_t\right)^{-1} \sum_{t=1}^{T} \sqrt{c_t} Y_t$$

Note that for any $h \in \mathbb{R}$, $\mathrm{P}(W_{T,k} - W_{T,k'} = h) = 0$. Along the subsequence, we also have by Lemma C.1 that

$$\sqrt{n_m}(\mu_{n_m,k} - \mu_{n_m,k'}) \to \Delta\mu_{k,k'} \in [-\infty, \infty].$$

Fix some t_0, π_{t_0}, w_{t_0} where $w_{t_0,k} - w_{t_0,k'} \neq -\Delta\mu_{kk'}$. It suffices to show that every sequence $t_m \to t_0, \pi_{t_m} \to \pi_{t_0}, w_{t_m} \to w_{t_0}$ in the support of U_m has that

$$\psi_n(t_m, \pi_{t_m}, w_{t_m}) \to (t_0, \{\mathbb{1}(w_{t_0,k} - w_{t_0,k'} + \Delta\mu_{kk'} > 0) : k, k' \in [K], k' \neq k\}, \{k : \pi_{t_m,k} = 0\}). \tag{C.9}$$

The convergence of t_m is immediate. Moreover, for all sufficiently large m, $t_m = t_0$. Since $w_{t_0,k} - w_{t_0,k'} \neq -\Delta\mu_{kk'}$, the indicator functions likewise converge. Lastly, note that if $\pi_{t_0,k} = 0$, then for all sufficiently large m, $\pi_{t_m,k} = 0$ as well since $\pi_{t_m,k} \notin (0, \epsilon)$ by Assumption C.3; likewise, if $\pi_{t_m,k} > \epsilon$, then for all sufficiently large m, $\pi_{t_m,k} > \epsilon$ as well. Hence $\{k : \pi_{t_m,k} = 0\} = \{k : \pi_{t_0,k} = 0\}$ for all sufficiently large m. This proves (C.9), and hence it proves the first part of Assumption C.4.

$\|\eta_n\| > \epsilon$ is by assumption. Lastly, note that for all sufficiently large m, $\lambda_{n_m} > 0$ and $\hat{\Sigma}_{n_m} > 0$. Hence, by Lemma C.9, we have that η_{n_m} is in the column space of

$$\frac{\lambda_{n_m}\lambda'_{n_m}/\sqrt{n_m}}{\lambda'_{n_m}1} + \frac{n_T}{n_m}\hat{\Sigma}_{n_m}^{-1}\hat{\Pi}_{Tn_m}$$

almost surely for all sufficiently large m. Similarly for the limit η. ☐

Lemma C.9. *Let $\mathbb{R}^K \ni u > 0$ be a vector with positive entries. Let $D \geq 0$ be a diagonal matrix with strictly positive entries on $\varnothing \neq J \subset [K]$. Consider $v \in \mathbb{R}^K$ whose nonzero entries are a subset of J. Then v is in the column space of $cuu' + D$ for $c > 0$.*

Proof. If $J = [K]$, then $cuu' + D$ is full rank. Thus v is in the column space. If $J \subsetneq [K]$, let $j_0 \in J^C$. Set $q \in \mathbb{R}^K$ where for $j \in J$, $q_j = v_j/D_{jj}$. Let $q_{j_0} = -\sum_{j\in J} u_j q_j$. Let $q_k = 0$ for $k \notin J \cup \{j_0\}$. By construction, $q'u = 0$ and $Dq = v$. This concludes the proof. ☐

Appendix D. Two-batch experiment and properties of Zhang et al. (2020)

To illustrate properties of Zhang et al. (2020)'s procedure, we restrict to a simple two-batch setup. Consider

$$X_1 \sim \mathcal{N}(\mu, 2I) \text{ and } X_2 \mid X_1 \sim \mathcal{N}(0, \operatorname{diag}(1/\Pi, 1/(1-\Pi))).$$

Since $\Pi_2 = (\Pi_{21}, 1-\Pi_{21})$ is summarized by Π_{21}, we drop the subscript and refer to $\Pi = \Pi_{21}$. Note that if the assignment algorithm is location-invariant—that is, $\Pi(X_1) = \Pi_t(X_1 + h1_2)$ for every h—then $\Pi_t(X_1) = \Pi_t(X_{11}, X_{12}) = \Pi_t(X_{11} - X_{12}, 0) = \Pi_t(X_{11} - X_{12})$ depends only on the difference $\Delta \equiv X_{11} - X_{12}$.

Suppose we are interested in making inferences about μ_1. Zhang et al. (2020) propose a simple inference procedure. The Gaussian limit analogue of their procedure is based on the following statistic, which is a sum of studentized batch means:

$$Z = \frac{1}{\sqrt{2}}X_{11} + \Pi(\Delta)^{1/2}X_{21}.$$

Since the studentized second batch statistic is independent of Δ,

$$\Pi^{1/2}(X_{21} - \mu_1) \mid \Delta \sim \mathcal{N}(0, 1),$$

the following quantity is pivotal:

$$Z^* = Z^*(\mu_1) = Z - \left(\frac{1}{\sqrt{2}} + \Pi^{1/2}\right)\mu_1 \sim \mathcal{N}(0, 2).$$

Thus, for example, we can test $H_0 : \mu_1 = \mu_{01}$ by comparing $Z^*(\mu_{01})$ to its distribution $\mathcal{N}(0, 2)$ under H_0. As another example, the natural estimator $\left(\frac{1}{\sqrt{2}} + \Pi^{1/2}\right)^{-1} Z$ is median-unbiased for μ_1.

Note that the joint density of X_1, X_2 factors into

$$p_\mu(X_1, X_2) = f_1(U\mu_1) f_2(V\mu_2) f_3(\Pi(\Delta)(\mu_1^2 - \mu_2^2)) h_1(\mu) h_2(X_1, X_2)$$

where $U = \frac{1}{2}X_{11} + \frac{1}{\Pi}X_{21}$, $V = \frac{1}{2}X_{12} + \frac{1}{1-\Pi}X_{22}$, h_1 is free of (X_1, X_2), and h_2 is free of μ. As a result, the statistics (U, V, Π) are sufficient for μ in the law of (X_1, X_2). Consequently, (U, V, Δ) is sufficient as well. We can decompose the statistic Z as a linear combination of the sufficient statistics (Δ, U, V) and a Gaussian noise term. We do so by studying the joint Gaussian distribution of (U, V, Z) conditional on Δ.

Proposition D.1. *The statistic Z obeys the following representation:*

$$\begin{aligned}
Z &= a\Delta + b(\Pi)V + c(\Pi)U + \sigma(\Pi)\xi \\
&= \frac{\Delta}{2\sqrt{2}} + \frac{\sqrt{2}\Pi - \sqrt{\Pi}}{1 + 4\Pi - 4\Pi^2}V + \frac{\sqrt{2} + 5\sqrt{\Pi} - \sqrt{2}\Pi - 4\Pi^{3/2}}{1 + 4\Pi - 4\Pi^2}U \\
&\quad + \sqrt{\frac{(1 + 2\Pi - 2\sqrt{2\Pi})(1 - \Pi)}{1 + 4\Pi - \Pi^2}} \cdot \xi
\end{aligned}$$

for some $\xi \sim \mathcal{N}(0, 1)$ independently of (U, V, Δ).

Proof. Observe that

$$\begin{bmatrix} \overline{X}_1 \\ X_{21} \\ X_{22} \end{bmatrix} \mid \underbrace{X_{11} - X_{12}}_{\Delta} \sim \mathcal{N}\left(\begin{bmatrix} \frac{\mu_1 + \mu_2}{2} \\ \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/\Pi & 0 \\ 0 & 0 & 1/(1-\Pi) \end{bmatrix} \right)$$

where $\overline{X}_1 = \frac{1}{2}(X_{11} + X_{12})$. Note that, since $X_{11} = \overline{X}_1 + \Delta/2$ and $X_{12} = \overline{X}_1 - \Delta/2$, the above display captures the distribution of the data conditional on Δ. Hence we can compute that

$$\begin{bmatrix} Z - \frac{\Delta}{2\sqrt{2}} \\ U - \frac{\Delta}{4} \\ V + \frac{\Delta}{4} \end{bmatrix} \mid \Delta \sim \mathcal{N}\left(\begin{bmatrix} \frac{\mu_1+\mu_2}{2\sqrt{2}} + \sqrt{\Pi}\mu_1 \\ \frac{\mu_1+\mu_2}{4} + \Pi\mu_1 \\ \frac{\mu_1+\mu_2}{4} + (1-\Pi)\mu_2 \end{bmatrix}, \begin{bmatrix} \frac{3}{2} & \frac{1}{2\sqrt{2}} + \sqrt{\Pi} & \frac{1}{2\sqrt{2}} \\ \frac{1}{2\sqrt{2}} + \sqrt{\Pi} & \frac{1}{4} + \Pi & \frac{1}{4} \\ \frac{1}{2\sqrt{2}} & \frac{1}{4} & \frac{5}{4} - \Pi \end{bmatrix} \right),$$

since the left-hand side are functions of $\overline{X}_1, X_{21}, X_{22}$ and Δ. The claim follows by computing that

$$\mathbb{E}[Z \mid \Delta, U, V] = a\Delta + b(\Pi)V + c(\Pi)U$$

and $\text{Var}(Z \mid \Delta, U, V) = \sigma^2(\Pi)$, since Z, U, V are jointly Gaussian conditional on Δ. □

One can further show that the projection of the pivot Z^* onto (Δ, V, U) is Gaussian with conditional mean $a(\Delta + \mu_2 - \mu_1)$ and conditional variance $\frac{3}{2} - \sigma^2(\Pi)$:

$$a\Delta + b(\Pi)V + c(\Pi)U - \left(\frac{1}{\sqrt{2}} + \Pi^{1/2}\right)\mu \mid \Delta \sim \mathcal{N}\left(a(\Delta + \mu_2 - \mu_1), \frac{3}{2} - \sigma^2(\Pi)\right)$$

As a result, the unconditional distribution of the projection of the pivot onto (Δ, U, V) is a mean and variance mixture of Gaussians, whose unconditional mean is zero. The pivot Z^* is thus Gaussian by injecting Δ-dependent noise $\sigma(\Pi)\xi$ so as to stabilize the Δ-conditional variance.

Due to the presence of the noise term ξ, the statistic Z suffers from certain deficiencies. For instance, there exist *recognizable subsets*—events on which inferences based on Z are unreliable.

Corollary D.1. *Assume that $\Pi \neq \frac{1}{2}$ with positive probability and $\Pi \in (0,1)$ almost surely. Fix any significance level $\alpha \in (0, 1/2)$. There exists an event A which does not depend on the unknown parameter μ on which the coverage probability is at most $1/2$. That is, for all $\mu \in \mathbb{R}^2$, $\mathrm{P}_\mu(A) > 0$, and*

$$\mathrm{P}_\mu\left(\left|Z - \left(\frac{1}{\sqrt{2}} + \Pi^{1/2}\right)\mu\right| \leq \sqrt{2}\Phi^{-1}(1-\alpha/2) \mid A\right) \leq \frac{1}{2}.$$

Proof. Let $A = \left\{|\xi| > \sqrt{2}\frac{1}{\sigma(\Pi)}\Phi^{-1}(1-\alpha/2), \Pi \neq 1/2\right\}$. Since when $\Pi \neq 1/2$, $\sigma(\Pi) > 0$, and $\Pi \neq 1/2$ with positive probability, A occurs with positive probability. Consider $Z = a\Delta + b(\Pi)V + c(\Pi)U + \sigma(\Pi)\xi$ and $Z = a\Delta + b(\Pi)V + c(\Pi)U - \sigma(\Pi)\xi$. Note that on the event A,

$$|Z - Z'| > 2\sqrt{2}\Phi^{-1}(1-\alpha/2).$$

Hence, on A,

$$|Z - (1/\sqrt{2} + \Pi^{1/2})\mu| < \sqrt{2}\Phi^{-1}(1-\alpha/2) \text{ and } |Z' - (1/\sqrt{2} + \Pi^{1/2})\mu| < \sqrt{2}\Phi^{-1}(1-\alpha/2)$$

are mutually exclusive events. Moreover, the probabilities of these two events conditional on A are identical by symmetry. Therefore,

$$2\,\mathrm{P}\left[|Z - (1/\sqrt{2} + \Pi^{1/2})\mu| < \sqrt{2}\Phi^{-1}(1-\alpha/2) \mid A\right] \leq 1.$$

This completes the proof. $\square$

Moreover, the natural estimator $\left(\frac{1}{\sqrt{2}} + \Pi^{1/2}\right)^{-1} Z$ can be improved in mean-squared error by Rao–Blackwellization against (U, V, Δ):

Corollary D.2. *Let $T_0 = \left(\frac{1}{\sqrt{2}} + \Pi^{1/2}\right)^{-1} Z$. Let*

$$T^* = \mathbb{E}\left[T_0 \mid U, V, \Delta\right] = \left(\frac{1}{\sqrt{2}} + \Pi^{1/2}\right)^{-1} (a\Delta + b(\Pi)V + c(\Pi)U)$$

be its conditional expectation with respect to U, V, Δ. Then, for any $\mu \in \mathbb{R}^2$,

$$\mathbb{E}_\mu[(T^* - \mu_1)^2] \leq \mathbb{E}_\mu[(T_0 - \mu_1)^2].$$

Proof. Note that we can write $T_0 = T^*(U, V, \Delta) + q(\Pi)\xi$ for some $q(\Pi)$. Note that

$$\mathbb{E}[(T_0 - \mu_1)^2] = q(\Pi)^2 + \mathbb{E}[(T^* - \mu_1)^2] \geq \mathbb{E}[(T^* - \mu_1)^2].$$

$\square$