Proofs and Supplementary Discussion

EC.1. Outline

This technical appendix is organized as follows.

1. Section EC.2 describes links with alternative performance measures.

2. Section EC.3 describes a numerical algorithm that can be used to implement TTPS.

3. Section EC.4 provides a more precise discussion of related work by Ryzhov (2016).

4. The theoretical analysis begins in Section EC.5. There we begin by noting some basic facts of exponential family distributions, as well as some results relating martingales to their quadratic variation process.

5. Section EC.6 establishes results related to the concentration of the posterior distribution, including the proofs of Prop. 4, Prop. 5, and Lemma 4.

6. Section EC.7 studies and simplifies the optimal exponents Γ^* and Γ^*_β, including the proofs of Lemma 2, Prop. 7, Lemma 3, Prop. 1, and Prop. 2.

7. We conclude with Section EC.8, which studies the top-two allocation rules and provides a proof of Prop. 8.

EC.2. Alternative Performance Criteria

By studying the rate of posterior convergence, over a class of priors satisfying Assumption 1, this paper essentially compares allocation rules in terms of the rate at which the evidence they gather would convince a Bayesian observer. This arguably allows for a more elegant analysis of top-two sampling algorithms, but extensions to more conventional performance criteria stands as an important open question. This section describes both close connections and important distinctions between the paper's analysis and alternative performance criteria studied in the literature.

Recall that, as a byproduct of our analysis, the paper characterizes the long-run fraction of samples the proposed top-two sampling algorithms collect from each arm on each problem instance $\boldsymbol{\theta}^*$. Namely, under the proposed top-two sampling algorithms with any parameter β, $\overline{\boldsymbol{\psi}}_n$ converges almost surely to $\boldsymbol{\psi}^{\beta}(\boldsymbol{\theta}^*)$, which attains the maximum in equation (12) defining the optimal constrainted exponent Γ^*_β. With appropriate tuning, measurement effort converges to $\boldsymbol{\psi}^*(\boldsymbol{\theta}^*)$, which attains the maximum in equation (11) defining the optimal exponent Γ^*.

As discussed in the literature review, the allocation $\boldsymbol{\psi}^*(\boldsymbol{\theta}^*)$ has been previously derived, seemingly independently, by Chernoff (1959) and Jennison et al. (1982). Nearly the same allocation was derived by Glynn and Juneja (2004) in the fixed-budget setting. These papers all focus on different performance criteria and so convergence to this allocation may be of interest even to readers who are skeptical of using the rate of posterior convergence as a comparator for adaptive algorithms. Unfortunately, there are also substantial subtleties since asymptotic convergence to this allocation may not ensure optimality according to these criteria. Below, we will give pointers to the literature on what is known, and not known, about different notions of optimality in best-arm identification. We will also discuss work of Garivier and Kaufmann (2016), which appeared simultaneously with an early version of this paper in COLT 2016 and a follow up paper by Qin et al. (2017), which was completed while this paper was under review.

Maximizing frequentist concentration of posterior beliefs: This paper benchmarks allocation rules in terms of the rate at which the evidence they gather would convince a rational Bayesian agent. More precisely, for any agent whose prior beliefs obey some regularity conditions stated in Assumption 1, we look at the rate their beliefs about which arm is optimal converge on the truth, under any fixed problem instance $\boldsymbol{\theta}^*$. It is shown that, the rate of convergence of such posterior beliefs is maximized when the algorithm samples arms with frequencies converging asymptotically to $\boldsymbol{\psi}^*(\boldsymbol{\theta}^*)$.

Minimizing frequentist probability of incorrect selection: The optimal fixed allocation $\boldsymbol{\psi}^*(\boldsymbol{\theta}^*)$ can also be derived from a fully frequentist perspective. Glynn and Juneja (2004) consider a model in which the decision-maker samples arms for n periods and then returns the arm $\hat{I}_n$ with highest observed mean. They suppose an agent samples arms non-adaptively, with fraction of effort allocated to each arm given by some strictly positive probability vector $\boldsymbol{\psi}$. They derive some choice of $\boldsymbol{\psi}$ (as a function of $\boldsymbol{\theta}$) that maximizes $\lim_{n\to\infty} -n^{-1}\log \mathbb{P}_{\boldsymbol{\theta}^*}\left(\hat{I}_n \neq \arg\max_i \theta_i\right)$, which is the exponent governing the rate at which the probability of incorrect selection decays. Formally, this $\boldsymbol{\psi}$ solves $\max_{\boldsymbol{\psi}} \min_{\boldsymbol{\theta}\in\Theta^c_{I^*}} D_{\psi}(\boldsymbol{\theta}||\boldsymbol{\theta}^*)$, which is the same as the optimal exponent in (7) except that the position of $\boldsymbol{\theta}^*$ and $\boldsymbol{\theta}$ is flipped. Unfortunately, the allocation solving this maximization problem cannot be implemented without knowledge of the arm means $\boldsymbol{\theta}^*$. Moreover, asymptotic convergence of sampling proportions to this limit does not guarantee the probability of incorrect selection decays at the optimal large deviations rate.

Ensuring a high probability of correct selection with minimal average sample size: Yet another fully frequentist problem formulation also leads to a derivation of the allocation $\boldsymbol{\psi}^*(\boldsymbol{\theta}^*)$. In this setup, the algorithm continues gathering samples adaptively, until a time τ at which it chooses to stop and return an estimate $\hat{I}_\tau$ of the optimal arm. The objective is to minimize the average number of samples collected while guaranteeing a probability of incorrect selection no greater than a specified level $\delta > 0$. Concurrently with this paper, Garivier and Kaufmann (2016) independently derived many of the same mathematical expressions while focusing on average sample size as a performance criterion. Precisely, they exhibit an algorithm whose average sample complexity scales as $\mathbb{E}_{\theta^*}[\tau_\delta] = \log(1/\delta)/\Gamma^* + o(\log(1/\delta))$ as $\delta \to 0$ and show this un-improvable. Their analysis also uncovers the fixed allocation $\boldsymbol{\psi}^*(\boldsymbol{\theta}^*)$ and their algorithm is specifically designed to attain asymptotically optimal bounds by converging on this allocation. Prior to the work of Garivier and Kaufmann (2016), it appears that Jennison et al. (1982) had attained similar results in a more restricted setting. A natural question is whether almost sure convergence to the allocation $\boldsymbol{\psi}^*(\boldsymbol{\theta}^*)$, as has been established here for top-two algorithms, is enough to ensure the algorithm requires minimal average sample size. This is investigated in follow up work by Qin et al. (2017), who show that one needs an additional condition ensuring that such convergence is not too slow, and establishes this property for a top-two sampling algorithm.

Minimizing exploration costs plus final regret: Chernoff (1959) considers much a broader set of sequential experimentation and hypothesis testing problems, albeit under much stronger technical restrictions than Garivier and Kaufmann (2016). He considers an experimenter who adaptively gathers information and then stops and returns an estimate of the true hypothesis. Specialized to the best-arm identification problem, his objective is to minimize $\mathbb{E}[c\tau + (\theta_{I^*} - \theta_{\hat{I}_\tau})]$ where c is a fixed cost of experimentation, τ is the stopping time, and $\hat{I}_\tau$ is the estimate of the best arm. For tractability, the paper studies optimal experimentation in the limit where the cost of experimentation c tends to 0. Specialized to this problem, the calculations in Chernoff (1959) essentially also give rise to the asymptotic proportions $\boldsymbol{\psi}^*(\boldsymbol{\theta}^*)$ and complexity term Γ^*. For ease of exposition, Chernoff (1959) makes numerous technical simplifications. In particular, he only considers problems with a finite parameter space. Interestingly, in this case Bayesian and frequentist solutions essentially coincide, in the sense that an algorithm that minimizes expected cost as $c \to 0$ on under problem instances drawn from a Bayesian prior should also minimize expected cost (among a reasonable class of algorithms) for every specific problem instance $\boldsymbol{\theta}^*$.

EC.3. An Implementation of TTPS

This section describes an implementation of the top-two probability sampling for a problem with a Beta prior and binary observations. In this problem, measurements are binary with success probability given by $\mathbb{P}(Y_{n,i} = 1) = \theta_i^*$. The algorithm begins with an independent prior, under which the ith component of $\boldsymbol{\theta}$ follows a Beta distribution with parameters $(\lambda_i^1, \lambda_i^2)$. When $\lambda_i^2 = \lambda_i^2 = 1$, this specifies a uniform prior over $[0,1]$. This prior distribution can be easily updated to form a posterior distribution according to the update rule given in line 19 of Algorithm 4.

This algorithm uses quadrature to approximate the integral defining $\alpha_{n,i}$. To understand this implementation, consider a random vector $(X_1, .., X_K)$ whose components are independently distributed with $X_i \sim \mathrm{Beta}(\lambda_i^1, \lambda_i^2)$. Then, the probability component i is maximal can be computed according to

$$
\begin{aligned}
\mathbb{P}(X_i = \max_j X_j) &= \int_{x\in\mathbb{R}} \mathbb{P}(\cap_{j\neq i}\{X_j \leq x\})\mathbb{P}(X_i = dx) \\
&= \int_{x\in\mathbb{R}} \left[\prod_{j\neq i} \mathbb{P}(X_j \leq x)\right] \mathbb{P}(X_i = dx)
\end{aligned}
$$

$$= \int_{x\in\mathbb{R}} \left[\left(\prod_{j=1}^{K} \mathbb{P}(X_j \le x) \right) / \mathbb{P}(X_i \le x) \right] \mathbb{P}(X_i = dx).$$

Algorithm 4 takes as input a vector of $\boldsymbol{x}$ consisting of M points in $(0,1)$ and approximates the above integral using quadrature at these points. The algorithm computes and updates the posterior PDF and CDF of θ_i in M dimensional vectors $\boldsymbol{f}_i$ and $\boldsymbol{F}_i$. It also stores and updates a vector $\overline{\boldsymbol{F}}$, where $\overline{F}_m = \prod_{i=1}^{K} F_{i,m}$ is the posterior probability all the designs have quality below x_m. Using these quantities, the posterior probability design i is optimal is approximated by a sum in line 8. Lines 11-15 select an action according to TTPS and lines 18-21 update the stored statistics of the posterior using Bayes rule. The algorithm continues for N time steps, and upon stopping returns the posterior parameters $\boldsymbol{\lambda}^1$ and $\boldsymbol{\lambda}^2$, which summarize all evidence gathered throughout the measurement process. The algorithm has $O(NKM)$ space and time complexity. It is worth noting that most operations in this algorithm can be implemented in a "vectorized" fashion in languages like MATLAB, NumPy, and Julia.

Algorithm 4 BernoulliTTPS($\beta, K, M, N, \boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2, \boldsymbol{x}$)

\\\\Initialize:

$f_{i,m} \leftarrow \text{Beta.pdf}(x_m|\lambda_i^1, \lambda_i^2) \qquad \forall i, m$

$F_{i,m} \leftarrow \text{Beta.cdf}(x_m|\lambda_i^1, \lambda_i^2) \qquad \forall i, m$

$\overline{F}_m \leftarrow \prod_i F_{i,m} \qquad \forall m$

for $n = 1 \ldots N$ **do**

  \\\\Compute Optimal Action Probabilities:

  $\alpha_i \leftarrow \sum_m f_{i,m} \overline{F}_m / F_{i,m} \qquad \forall i$

  \\\\Act and Observe:

  $J_1 \leftarrow \arg\max_i \alpha_i$

  $J_2 \leftarrow \arg\max_{i \neq J_1} \alpha_i$

  Sample $B \sim \text{Bernoulli}(\beta)$

  $I \leftarrow BJ_1 + (1-B)J_2$.

  Play I and Observe $Y_{n,I} \in \{0,1\}$.

  \\\\Update Statistics:

  $(\lambda_I^1, \lambda_I^2) \leftarrow (\lambda_I^1, \lambda_I^2) + (Y_{n,I}, 1 - Y_{n,I})$

  $\overline{F}_m \leftarrow (\overline{F}_m / F_{I,m}) \times \text{Beta.cdf}(x_m|\lambda_I^1, \lambda_I^2) \qquad \forall m$

  $F_{I,m} \leftarrow \text{Beta.cdf}(x_m|\lambda_I^1, \lambda_I^2) \qquad \forall m$

  $f_{I,m} \leftarrow \text{Beta.pdf}(x_m|\lambda_I^1, \lambda_I^2) \qquad \forall m$

end for

return $\boldsymbol{V}, \boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2$

EC.4. Discussion of the Expected Improvement Algorithm

Here, we briefly discuss interesting recent results of Ryzhov (2016). He studies a setting with an uncorrelated Gaussian prior, and Gaussian observation noise $Y_{n,i} \sim N(\theta_i, \sigma_i^2)$. To simplify our discussion, let us restrict attention to the case of common variance $\sigma_1 = \cdots = \sigma_k = \sigma$. Ryzhov (2016) shows that under the the expected-improvement algorithm, in the limit as $n \to \infty$

$$\sum_{i \neq I^*} \Psi_{n,i} = O(\log n) \tag{EC.1}$$

and

$$\Psi_{n,i}(\theta_I^* - \theta_i)^2 \sim \Psi_{n,j}(\theta_I^* - \theta_j)^2 \qquad \forall i, j \neq I^* \tag{EC.2}$$

Recall that $\Psi_{n,i} = \sum_{\ell=1}^{n} \psi_{n,i}$ denotes the total measurement effort allocated to design i. The sampling ratios (EC.2) are the ratios suggested in the optimal computing budget allocation of Chen et al. (2000). This work therefore establishes an interesting link between EI and OCBA, which appear quite different on the surface.

Unfortunately, property (EC.1) is *not suggested* by the OCBA, and implies that $\Pi_n(\Theta_{I^*}^c)$ cannot tend to zero at an exponential rate. To see this precisely, assume without loss of generality that $I^* = 1$. Then (EC.1) implies $\overline{\boldsymbol{\psi}}_n \to \mathbf{e}_1 \equiv (1, 0, 0, ..., 0)$. It is easy to show that $\min_{\boldsymbol{\theta} \in \Theta_1^c} D_{\mathbf{e}_1}(\boldsymbol{\theta}^* || \boldsymbol{\theta}) = 0$ and therefore, by Proposition 5,

$$\lim_{n \to \infty} -n^{-1} \log \Pi_n(\Theta_1^c) = 0.$$

It is also worth noting that the sampling ratios in (EC.2) are not actually optimal for any finite number of designs k. Specifying our calculations as in Example 1, one can show that under an optimal fixed allocation $(\psi_i, ..., \psi_k)$,

$$\frac{(\theta_{I^*}^* - \theta_i^*)^2}{1/\psi_{I^*} + 1/\psi_i} = \frac{(\theta_{I^*}^* - \theta_j^*)^2}{1/\psi_{I^*} + 1/\psi_j} \qquad \forall i, j \neq I^*.$$

These calculations match those in Glynn and Juneja (2004) and Jennison et al. (1982). As a result, there is no problem with finite k for which the sampling ratios in (EC.2) are optimal5. One can show, in fact, that any optimal multi-armed bandit algorithm that attains the lower bound of Lai and Robbins (1985) also satisfies equations (EC.1) and (EC.2). The main innovation in this paper is to show how to build on such bandit algorithms to attain near-optimal rates for the best-arm identification problem.

Ryzhov (2016) also studies the knowledge gradient policy, which could offer improved performance as (EC.1) no longer holds, but shows that as $n \to \infty$

$$\Psi_{n,i}(\theta_I^* - \theta_i) \sim \Psi_{n,j}(\theta_I^* - \theta_j) \qquad \forall i,j \neq I^*,$$

which could be very far from the optimal sampling proportions.

EC.5. Preliminaries

This section presents some basic results which will be used in the subsequent analysis. First, unless clearly specified, all statements about random variables are meant to hold with probability 1. So for sequences of random variables $\{X_n\}$ and $\{Y_n\}$, if we say that $X_n \to \infty$ whenever $Y_n \to \infty$, this means that the set $\{\omega : Y_n(\omega) \to \infty, X_n(\omega) \nrightarrow \infty\}$ has measure zero.

Facts about the exponential family. The log partition function $A(\theta)$ is strictly convex and differentiable, with

$$A'(\theta) = \int T(y) p(y|\theta) d\nu(y) \tag{EC.3}$$

equal to the mean under θ. The Kullback-Leibler divergence is equal to

$$d(\theta||\theta') = (\theta - \theta')A'(\theta) - A(\theta) + A(\theta') \tag{EC.4}$$

and satisfies

$$\theta'' > \theta' \geq \theta \implies d(\theta||\theta'') > d(\theta||\theta') \tag{EC.5}$$

$$\theta'' > \theta' \leq \theta \implies d(\theta||\theta'') < d(\theta||\theta'). \tag{EC.6}$$

Finally, since $[\underline{\theta}, \overline{\theta}]$ is bounded, and we have assumed $\sup_{\theta \in [\underline{\theta}, \overline{\theta}]} |A'(\theta)| < \infty$,

$$\sup_{\theta \in [\underline{\theta}, \overline{\theta}]} |A(\theta)| < \infty \quad \text{and} \quad \sup_{\theta, \theta' \in [\underline{\theta}, \overline{\theta}]} d(\theta||\theta') < \infty. \tag{EC.7}$$

This effectively guarantees no single observation can provide enough information to completely rule out a parameter.

Some martingale convergence results. The next fact relates the behavior of a martingale M_n to its quadratic variation $\langle M \rangle_n$.

LEMMA EC.1. *(Williams (1991), 12.13-12.14) Let $\{M_n\}$ be a square-integrable martingale adapted to the filtration $\{\mathcal{H}_n\}$ and let*

$$\langle M \rangle_n = \sum_{\ell=1}^{n} \mathbb{E}[(M_\ell - M_{\ell-1})^2 \,|\mathcal{H}_{\ell-1}]$$

denote the corresponding quadratic variation process. Then

$$\frac{M_n}{\langle M \rangle_n} \to \infty$$

almost surely if $\langle M \rangle_n \to \infty$ and $\lim_{n\to\infty} M_n$ exists and is finite almost surely if $\lim_{n\to\infty} \langle M \rangle_n < \infty$.

The next lemma is crucial to our analysis. To draw the connection with our setting, imagine an adaptive-randomized rule is used to determine when to draw samples from a population. Here $Y_n \in \mathbb{R}$ denotes the sample at time n, $X_n \in \{0,1\}$ indicates whether the sample was measured, and $Z_n \in [0,1]$ determines the probability of measurement conditioned on the past. This lemma provides a law of large numbers when measurement effort $\sum_{\ell=1}^{n} Z_\ell$ tends to infinity, but shows that if measurement effort is finite then $\sum_{\ell=1}^{\infty} X_\ell Y_\ell$ is also finite; in this sense the observations collected from Y_n are inconclusive when measurement effort is finite.

LEMMA EC.2. *Let $\{Y_n\}$ be an i.i.d sequence of real-valued random variables with finite variance and let $\{X_n\}$ be a sequence of binary random variables. Suppose each sequence is adapted to the filtration $\{\mathcal{H}_n\}$, and define $Z_n = \mathbb{P}(X_n = 1|\mathcal{H}_{n-1})$. If Y_n is independent of $(\mathcal{H}_{n-1}, X_n)$, then with probability 1,*

$$\lim_{n\to\infty} \sum_{\ell=1}^{n} Z_\ell = \infty \implies \lim_{n\to\infty} \frac{\sum_{\ell=1}^{n} X_\ell Y_\ell}{\sum_{\ell=1}^{n} Z_\ell} = \mathbb{E}[Y_1]$$

and

$$\lim_{n\to\infty} \sum_{\ell=1}^{n} Z_\ell < \infty \implies \sup_{n\in\mathbb{N}} \left| \sum_{\ell=1}^{n} X_\ell Y_\ell \right| < \infty.$$

Proof. Let $\mu = \mathbb{E}[Y_1]$ and $\sigma^2 = \mathbb{E}[(Y_1 - \mathbb{E}[Y_1])^2]$ denote the mean and variance of each Y_n. Define the martingale

$$M_n = \sum_{\ell=1}^{n} (X_\ell Y_\ell - Z_\ell \mu)$$

with $M_0 = 0$. Put $S_n = \sum_{\ell=1}^{n} Z_\ell$. This martingale has quadratic variation

$$\begin{aligned}
\langle M \rangle_n &= \sum_{\ell=1}^{n} \mathbb{E}[(M_\ell - M_{\ell-1})^2 | \mathcal{H}_{\ell-1}] \\
&= \sum_{\ell=1}^{n} \mathbb{E}[(X_\ell(Y_\ell - \mu) + (X_\ell - Z_\ell)\mu)^2 \,|\mathcal{H}_{\ell-1}] \\
&= \sum_{\ell=1}^{n} Z_\ell \sigma^2 + \sum_{\ell=1}^{n} Z_\ell (1 - Z_\ell)\mu^2 \qquad (\text{since } \mathbb{E}[(Y_\ell - \mu)(X_\ell - Z_\ell) \mid \mathcal{H}_{\ell-1}] = 0) \\
&\leq (\sigma^2 + \mu^2) S_n.
\end{aligned}$$

We use the shorthand $S_\infty = \lim_{n\to\infty} S_n$ and $\langle M \rangle_\infty = \lim_{n\to\infty} \langle M \rangle_n$.

Suppose $S_\infty < \infty$ so $\langle M \rangle_\infty < \infty$. By Lemma EC.1, $\lim_{n\to\infty} M_n$ exists and is finite almost surely, which implies $\sup_{n\in\mathbb{N}} |M_n| < \infty$. Since $|\sum_{\ell=1}^{n} X_\ell Y_\ell| = |M_n - \mu S_n| \leq |M_n| + |\mu S_\infty|$, this shows $\sup_{n\in\mathbb{N}} |\sum_{\ell=1}^{n} X_\ell Y_\ell| < \infty$ as desired.

Now, suppose $S_\infty = \infty$. If $\langle M \rangle_\infty < \infty$, then again by Lemma EC.1, $\lim_{n\to\infty} M_n < \infty$ and it is immediate that $S_n^{-1} M_n \to 0$. However, if $\langle M \rangle_\infty = \infty$ then

$$\frac{M_n}{\langle M \rangle_n} \to 0,$$

which implies $S_n^{-1} M_n \to 0$ since $S_n \geq (\sigma^2 + \mu^2)\langle M \rangle_n$. $\square$

Taking $Y_n = 1$ in the lemma above yields Levy's extension of the Borel–Cantelli lemmas (Williams (1991), 12.15). Specialized to our setting, this result relates the long run measurement effort $\Psi_{n,i} = \sum_{\ell=1}^{n} \psi_{n,i}$ to the number of times alternative i is actually measured $\sum_{\ell=1}^{n} \mathbf{1}(I_n = i)$.

Corollary EC.1. *For $i \in \{1, ..., k\}$, set $S_{n,i} = \sum_{\ell=1}^{n} \mathbf{1}(I_n = i)$. Then, with probability 1,*

$$\Psi_{n,i} \to \infty \iff S_{n,i} \to \infty$$

and

$$\Psi_{n,i} \to \infty \implies \frac{S_{n,i}}{\Psi_{n,i}} \to 1.$$

Proof. Apply Lemma EC.2 with $Y_n = 1$, $X_n = \mathbf{1}(I_n = 1)$, and $\mathcal{H}_n = \mathcal{F}_n$. (Recall $\mathcal{F}_n$ denotes the sigma algebra generated by $(I_1, Y_{1,I_1}, ... I_n, Y_{n,I_n})$.) Then $Z_n = \psi_{n,i}$ by definition. $\square$

EC.6. Posterior Concentration and anti-Concentration

EC.6.1. Uniform Convergence of the Log-Likelihood

We study the log-likelihood

$$\Lambda_n(\boldsymbol{\theta}^*||\boldsymbol{\theta}) \triangleq \log\left(\frac{L_n(\boldsymbol{\theta}^*)}{L_n(\boldsymbol{\theta})}\right) = \sum_{\ell=1}^{n} \log\left(\frac{p(Y_{\ell,I_\ell}|\theta^*_{I_\ell})}{p(Y_{\ell,I_\ell}|\theta_{I_\ell})}\right) \tag{EC.8}$$

and the log-likelihood from observations of design i

$$\Lambda_{n,i}(\theta_i^*||\theta_i) \triangleq \sum_{\ell=1}^{n} \mathbf{1}(I_n = i) \log\left(\frac{p(Y_{n,i}|\theta_i^*)}{p(Y_{n,i}|\theta_i)}\right).$$

A Doob-decomposition expresses $\Lambda_{n,i}(\theta_i) = A_n(\theta_i) + M_n(\theta_i)$ as the sum of an $\mathcal{F}_{n-1}$ predictable process $A_n(\theta_i)$ and a Martingale $M_n(\theta_i)$. Moreover, an easy calculation shows $A_n(\theta_i) = \Psi_{n,i} d(\theta_i^*||\theta_i)$ and $M_n(\theta_i) = \Lambda_{n,i}(\theta_i^*||\theta_i) - \Psi_{n,i} d(\theta_i^*||\theta_i)$. Applying Lemma EC.2 shows $\Psi_{n,i}^{-1} M_n(\theta_i) \to 0$ if $\Psi_{n,i} \to \infty$, which shows the log-likelihood ratio tends to infinity at rate $\Psi_{n,i} d(\theta_i^*||\theta_i)$. The next lemma strengthens this, and provides a link between these quantities that holds uniformly in θ_i.

Lemma EC.3. *With probability 1, if $\Psi_{n,i} \to \infty$ as $n \to \infty$ then*

$$\sup_{\theta_i \in [\underline{\theta}, \overline{\theta}]} \Psi_{n,i}^{-1} |\Lambda_{n,i}(\theta_i^*||\theta_i) - \Psi_{n,i} d(\theta_i^*||\theta_i)| \to 0,$$

and if $\lim_{n\to\infty} \Psi_{n,i} < \infty$ *then*

$$\sup_{\theta_i \in [\underline{\theta}, \overline{\theta}]} \sup_{n \in \mathbb{N}} |\Lambda_{n,i}(\theta_i)| + |\Psi_{n,i} d(\theta_i^*||\theta_i)| < \infty.$$

Proof. Define $\xi_n \triangleq T(Y_{n,i}) - \mathbb{E}[T(Y_{n,i})]$ and $X_n \triangleq \mathbf{1}(I_n = i)$. Note that $\mathbb{E}[\xi_n|\mathcal{F}_{n-1}] = 0$, $\mathbb{E}[X_n|\mathcal{F}_{n-1}] = \psi_{n,i}$, and, conditioned on $\mathcal{F}_{n-1}$, X_n is independent of ξ_n. Using the form of the exponential family density given in equation (1), and the form of the KL-divergence given in equation (EC.4), the log-likelihood ratio can be written as

$$\begin{aligned}
\log\left(\frac{p(Y_{n,i}|\theta_i^*)}{p(Y_{n,i}|\theta_i)}\right) &= (\theta_i^* - \theta_i) T(Y_{n,i}) - (A(\theta_i^*) - A(\theta_i)) \\
&= d(\theta_i^*||\theta_i) + (\theta_i^* - \theta_i)\left(T(Y_{n,i}) - \mathbb{E}[T(Y_{n,i})]\right) \\
&= d(\theta_i^*||\theta_i) + (\theta_i^* - \theta_i)\xi_n
\end{aligned}$$

Therefore,

$$\begin{aligned}\Lambda_{n,i}(\theta_i^*||\theta_i) - \Psi_{n,i} d(\theta_i^*||\theta_i) &= \sum_{\ell=1}^{n} X_\ell \log\left(\frac{p(Y_{\ell,i}|\theta_i^*)}{p(Y_{\ell,i}|\theta_i)}\right) - \sum_{\ell=1}^{n}\psi_{\ell,i} d(\theta_i^*||\theta_i)\\ &= \sum_{\ell=1}^{n}(X_\ell - \psi_{\ell,i}) d(\theta_i^*||\theta_i) + \sum_{\ell=1}^{n} X_\ell \xi_\ell(\theta_i^* - \theta_i).\end{aligned}$$

Here $|\theta_i^* - \theta_i| \leq \overline{\theta} - \underline{\theta} \equiv C_2$ is bounded uniformly. Similarly, as shown in Appendix EC.5, $d(\theta_i^*||\theta_i)$ is bounded uniformly in θ_i by

$$C_1 \equiv \max_{\theta' \in [\underline{\theta}, \overline{\theta}]} d(\theta_i^*||\theta_i') < \infty.$$

This implies,

$$|\Lambda_{n,i}(\theta_i) - \Psi_{n,i} d(\theta_i^*||\theta_i)| \leq C_1 \left|\sum_{\ell=1}^{n}(X_\ell - \psi_{\ell,i})\right| + C_2\left|\sum_{\ell=1}^{n} X_\ell\xi_\ell\right| \tag{EC.9}$$

$$|\Lambda_{n,i}(\theta_i)| \leq C_1\Psi_{n,i} + C_1 \left|\sum_{\ell=1}^{n}(X_\ell - \psi_{\ell,i})\right| + C_2\left|\sum_{\ell=1}^{n} X_\ell\xi_\ell\right|. \tag{EC.10}$$

Since $\mathbb{E}[\xi_n^2] < \infty$, the result then follows by applying Lemma EC.2 and Corollary EC.1. In particular, when $\Psi_{n,i} \to \infty$,

$$\lim_{n\to\infty} \Psi_{n,i}^{-1} \sum_{\ell=1}^{n}(X_\ell - \psi_{\ell,i}) = 0 \quad \text{and} \quad \lim_{n\to\infty} \Psi_{n,i}^{-1}\sum_{\ell=1}^{n} X_\ell \xi_\ell = 0$$

When $\lim_{n\to\infty}\Psi_{n,i} < \infty$,

$$\sup_{n\in\mathbb{N}} \left|\sum_{\ell=1}^{n}(X_\ell - \psi_{\ell,i})\right| < \infty \quad \text{and} \quad \sup_{n\in\mathbb{N}}\left|\sum_{\ell=1}^{n} X_\ell\xi_\ell\right| < \infty.$$

It is also immediate that $d(\theta_i^*||\theta_i)\Psi_{n,i} \leq C_1\Psi_{n,i} \nrightarrow \infty$, which by (EC.10) implies the second part of the result. □

A corollary of the previous lemma relates the log-likelihood ratio $\Lambda_n(\boldsymbol{\theta}^*||\boldsymbol{\theta})$ to the Kullback-Leibler divergence $D_{\overline{\psi}_n}(\boldsymbol{\theta}^*||\boldsymbol{\theta})$. Recall that $D_{\boldsymbol{\psi}}(\boldsymbol{\theta}^* \,||\, \boldsymbol{\theta}) = \sum_{i=1}^{k}\psi_i d(\theta_i^* \,||\, \theta_i)$ denotes a KL-divergence weighted under the probability vector $\boldsymbol{\psi}$ and $\overline{\psi}_{n,i} = \frac{\Psi_{n,i}}{n}$ denotes a kind of average measurement effort allocated to arm i.

Corollary EC.2. *With probability 1,*

$$\sup_{\boldsymbol{\theta}\in\Theta} |n^{-1}\Lambda_n(\boldsymbol{\theta}^*||\boldsymbol{\theta}) - D_{\overline{\boldsymbol{\psi}}_n}(\boldsymbol{\theta}^*||\boldsymbol{\theta})| \to 0$$

Proof.

$$\begin{aligned}\left|n^{-1}\Lambda_n(\boldsymbol{\theta}^*||\boldsymbol{\theta}) - D_{\overline{\psi}_n}(\boldsymbol{\theta}^*||\boldsymbol{\theta})\right| &= \left|n^{-1}\sum_{i=1}^{k}\left(\Lambda_{n,i}(\theta_i^*||\theta_i) - \Psi_{n,i}d(\theta_i^*||\theta_i)\right)\right| \\ &\leq \sum_{i=1}^{k} n^{-1}|\Lambda_{n,i}(\theta_i^*||\theta_i) - \Psi_{n,i}d(\theta_i^*||\theta_i)|.\end{aligned}$$

Lemma EC.3 implies

$$\sup_{\theta_i\in[\underline{\theta},\overline{\theta}]} n^{-1}|\Lambda_{n,i}(\theta_i^*||\theta_i) - \Psi_{n,i}d(\theta_i^*||\theta_i)| \to 0,$$

which completes the proof. $\square$

EC.6.2. Posterior Consistency: Proof of Prop. 4

PROPOSITION 4. *For any $i\in\{1,..,k\}$ if $\Psi_{n,i}\to\infty$, then, for all $\epsilon>0$*

$$\Pi_n(\{\boldsymbol{\theta}\in\Theta|\theta_i\notin(\theta_i^*-\epsilon,\theta_i^*+\epsilon)\})\to 0,$$

with probability 1. If $\mathcal{I}=\{i\in\{1,...,k\}|\lim_{n\to\infty}\Psi_{n,i}<\infty\}$ is nonempty, then

$$\inf_{n\in\mathbb{N}} \Pi_n(\{\boldsymbol{\theta}\in\Theta|\theta_i\in(\theta_i',\theta_i'')\ \forall i\in\mathcal{I}\})>0$$

for any collections of open intervals $(\theta_i',\theta_i'')\subset(\underline{\theta},\overline{\theta})$ ranging over $i\in\mathcal{I}$.

Because we don't assume an independent prior across the designs, Π_1 is not a product measure and therefore neither is Π_n. This makes it challenging to reason about the marginal posterior of each design, which is required for Proposition 4. Thankfully, since the prior density is bounded, Π_n behaves like a product measure. Note that the likelihood function can be written as the product of k terms:

$$L_n(\boldsymbol{\theta})=\prod_{i=1}^{k}L_{n,i}(\theta_i)$$

where

$$L_{n,i}(\theta_i)\triangleq\prod_{\substack{\ell\leq n\\ I_\ell=i}}p(Y_{\ell,1}|\theta_i)$$

with the convention that $L_{n,i}(\theta_i)=1$ when $\sum_{\ell=1}^{n}\mathbf{1}(I_\ell=i)=0$. Therefore $L_n(\boldsymbol{\theta})$ forms the density of a product measure. By normalizing, this induces a probability measure over Θ,

$$\mathcal{L}_n(\tilde{\Theta}) \triangleq \frac{\int_{\tilde{\Theta}} L_n(\boldsymbol{\theta})d\boldsymbol{\theta}}{\int_{\Theta} L_n(\boldsymbol{\theta})d\boldsymbol{\theta}} \qquad \tilde{\Theta}\subset\Theta,$$

which, as we argue in the next lemma, behaves like the posterior Π_n.

LEMMA EC.4. *For any set* $\tilde{\Theta}\subset\Theta$,

$$C^{-1}\mathcal{L}_n(\tilde{\Theta}) \le \Pi_{n+1}(\tilde{\Theta}) \le C\mathcal{L}_n(\tilde{\Theta}),$$

where

$$C=\frac{\sup_{\boldsymbol{\theta}\in\Theta}\pi_1(\boldsymbol{\theta})}{\inf_{\boldsymbol{\theta}\in\Theta}\pi_1(\boldsymbol{\theta})}<\infty$$

is independent of n *and* $\tilde{\Theta}$.

Proof. This follows immediately from Assumption 1 by bounding the prior density $\pi_1(\boldsymbol{\theta})$ from above and below in the relation

$$\Pi_{n+1}(\tilde{\Theta})=\frac{\int_{\tilde{\Theta}}\pi_1(\boldsymbol{\theta})L_n(\boldsymbol{\theta})d\boldsymbol{\theta}}{\int_{\Theta}\pi_1(\boldsymbol{\theta})L_n(\boldsymbol{\theta})d\boldsymbol{\theta}}.$$

□

We can now prove Proposition 4.

Proof of Proposition 4. We begin with the first part of the result. For simplicity of notation, we focus on the upper interval $\tilde{\Theta}=\{\boldsymbol{\theta}\in\Theta:\theta_i>\theta_i^*+\epsilon\}$, but results follow identically for the lower interval. We want to show $\Pi_n(\tilde{\Theta})\to 0$, which occurs if and only if $\mathcal{L}_n(\tilde{\Theta})\to 0$. Since $\mathcal{L}_n$ is a product measure,

$$\mathcal{L}_n(\tilde{\Theta})=\frac{\int_{\theta_i^*+\epsilon}^{\overline{\theta}} L_{n,i}(\theta)d\theta}{\int_{\underline{\theta}}^{\overline{\theta}} L_{n,i}(\theta)d\theta}=\frac{\int_{\theta_i^*+\epsilon}^{\overline{\theta}}\left(L_{n,i}(\theta)/L_{n,i}(\theta_i^*)\right)d\theta}{\int_{\underline{\theta}}^{\overline{\theta}}\left(L_{n,i}(\theta)/L_{n,i}(\theta_i^*)\right)d\theta}=\frac{\int_{\theta_i^*+\epsilon}^{\overline{\theta}}\exp\{-\Lambda_{n,i}(\theta_i^*||\theta)\}d\theta}{\int_{\underline{\theta}}^{\overline{\theta}}\exp\{-\Lambda_{n,i}(\theta_i^*||\theta)\}d\theta} \tag{EC.11}$$

where $\Lambda_{n,i}(\theta_i^*||\theta_i)=\log(L_{n,i}(\theta_i^*)/L_{n,i}(\theta_i)$. By Lemma EC.3, with probability 1 there is a sequence $a_n\to 0$ such that $|\Lambda_{n,i}(\theta_i^*||\theta)-\Psi_{n,i}d(\theta_i^*||\theta)|\le a_n$ for all θ. Then, for $b_n=e^{a_n}/e^{-a_n}\to 1$, one has

$$\mathcal{L}_n(\tilde{\Theta})\le\frac{b_n\int_{\theta_i^*+\epsilon}^{\overline{\theta}}\exp\{-\Psi_{n,i}d(\theta_i^*||\theta)\}d\theta}{\int_{\underline{\theta}}^{\overline{\theta}}\exp\{-\Psi_{n,i}d(\theta_i^*||\theta)\}d\theta}\le\frac{b_n\int_{\theta_i^*+\epsilon}^{\overline{\theta}}\exp\{-\Psi_{n,i}d(\theta_i^*||\theta)\}d\theta}{\int_{\theta_i^*}^{\theta_i^*+\epsilon/2}\exp\{-\Psi_{n,i}d(\theta_i^*||\theta)\}d\theta}.$$

The integral in the numerator is upper bounded by $(\overline{\theta}-\theta_i^*-\epsilon)\exp\{-\Psi_{n,i}d(\theta_i^*||\theta_i^*+\epsilon)\}$ while the integral in the denominator is lower bounded by $(\epsilon/2)\exp\{-\Psi_{n,i}d(\theta_i^*||\theta_i^*+\epsilon/2)\}$. This shows, if $\Psi_{n,i}\to\infty$ then

$$\mathcal{L}_n(\tilde{\Theta}) \le c_0 b_n \exp\{-\Psi_{n,i}\left(d(\theta_i^*||\theta_i^*+\epsilon)-d(\theta_i^*||\theta_i^*+\epsilon/2)\right)\}\to 0$$

where $c_0 = 2\epsilon^{-1}(\overline{\theta}-\theta_i^*-\epsilon)$.

The second part of the claim follows from the lower bound in Lemma EC.4 of

$$\Pi_{n+1}(\{\boldsymbol{\theta}\in\Theta|\theta_i\in(\theta',\theta'')\ \forall i\in\mathcal{I}\}) \ge C^{-1}\mathcal{L}_n(\{\boldsymbol{\theta}\in\Theta|\theta_i\in(\theta_i',\theta_i'')\ \forall i\in\mathcal{I}\}) \tag{EC.12}$$

$$= C^{-1}\prod_{i\in\mathcal{I}}\mathcal{L}_n(\{\boldsymbol{\theta}\in\Theta|\theta_i\in(\theta_i',\theta_i'')\}). \tag{EC.13}$$

As in (EC.11),

$$\mathcal{L}_n(\{\boldsymbol{\theta}\in\Theta|\theta_i\in(\theta_i',\theta_i'')\}) = \frac{\int_{\theta'}^{\theta''}\exp\{-\Lambda_{n,i}(\theta_i^*||\theta)\}d\theta}{\int_{\underline{\theta}}^{\overline{\theta}}\exp\{-\Lambda_{n,i}(\theta_i^*||\theta)\}d\theta}.$$

When $\lim_{n\to\infty}\Psi_{n,i}<\infty$, Lemma EC.3 shows that for each $i\in\mathcal{I}$,

$$\sup_{\theta_i\in[\underline{\theta},\overline{\theta}]}\sup_{n\in\mathbb{N}}|\Lambda_{n,i}(\theta_i^*||\theta_i)|<\infty.$$

This implies

$$\inf_n \mathcal{L}_n(\{\boldsymbol{\theta}\in\Theta|\theta_i\in(\theta_i',\theta_i'')\})>0$$

and establishes the claim. □

EC.6.3. Large Deviations: Proof of Proposition 5

All statements in this section hold when observations are drawn under the parameter $\boldsymbol{\theta}^*$. Since $\boldsymbol{\theta}^*$ is fixed throughout, we simplify notation and write

$$W_n(\boldsymbol{\theta}) \triangleq D_{\overline{\boldsymbol{\psi}}_n}(\boldsymbol{\theta}^*||\boldsymbol{\theta}).$$

Proposition 5 can be restated with notation as follows.

Proposition 5. *For any open set $\tilde{\Theta}\subset\Theta$,*

$$\Pi_n(\tilde{\Theta}) \doteq \exp\left\{-n\inf_{\boldsymbol{\theta}\in\tilde{\Theta}}W_n(\boldsymbol{\theta})\right\}.$$

Note that $W_n(\boldsymbol{\theta}^*)=0$. As shown in the next lemma $n^{-1}\log(\pi_n(\boldsymbol{\theta})/\pi_n(\boldsymbol{\theta}^*)) - W_n(\boldsymbol{\theta}) \to 0$ uniformly in $\boldsymbol{\theta}$.

LEMMA EC.5. *With probability 1,*

$$\sup_{\boldsymbol{\theta}\in\Theta} n^{-1}\left|\log\left(\frac{\pi_n(\boldsymbol{\theta}^*)}{\pi_n(\boldsymbol{\theta})}\right) - W_n(\boldsymbol{\theta})\right| \to 0.$$

Proof. Using the notation for the log-likelihood in equation (EC.8), we have

$$\log\left(\frac{\pi_n(\boldsymbol{\theta}^*)}{\pi_n(\boldsymbol{\theta})}\right) - W_n(\boldsymbol{\theta}) = \log\left(\frac{\pi_1(\boldsymbol{\theta}^*)}{\pi_1(\boldsymbol{\theta})}\right) + (\Lambda_{n-1}(\boldsymbol{\theta}^*||\boldsymbol{\theta}) - W_{n-1}(\boldsymbol{\theta})) + (W_{n-1}(\boldsymbol{\theta}) - W_n(\boldsymbol{\theta}))\,.$$

Since $\inf_{\boldsymbol{\theta}\in\Theta}\pi_1(\boldsymbol{\theta}) > 0$ and $\sup_{\boldsymbol{\theta}\in\Theta}\pi_1(\boldsymbol{\theta}) < \infty$, $n^{-1}\log(\pi_1(\boldsymbol{\theta})/\pi_1(\boldsymbol{\theta}^*)) \to 0$ uniformly in $\boldsymbol{\theta}$. By Corollary EC.2, $n^{-1}(\Lambda_{n-1}(\boldsymbol{\theta}) - W_{n-1}(\boldsymbol{\theta})) \to 0$ uniformly as well. Finally, by equation (EC.7), $n^{-1}(W_n(\boldsymbol{\theta}) - W_{n-1}(\boldsymbol{\theta})) \le n^{-1}\max_i d(\theta_i^*||\theta_i) \to 0$ uniformly in $\boldsymbol{\theta}$.

□

The remaining proof of Proposition 5 follows from a sequence of lemmas. The next observes a form of uniform continuity of W_n that follows from the uniform bound on $A'(\theta)$ in Assumption 1.

LEMMA EC.6. *For all $\epsilon>0$, there exists $\delta>0$ such that for $\boldsymbol{\theta},\boldsymbol{\theta}'\in\Theta$*

$$\|\boldsymbol{\theta}-\boldsymbol{\theta}'\|_\infty \le \delta \implies \sup_{n\in\mathbb{N}}|W_n(\boldsymbol{\theta}) - W_n(\boldsymbol{\theta}')| \le \epsilon.$$

Proof. Recall the formula for KL-divergence in exponential family distributions described in Appendix EC.5. We have that

$$\begin{aligned}
|W_n(\boldsymbol{\theta}) - W_n(\boldsymbol{\theta}')| &= \left|\sum_{i=1}^k \overline{\psi}_{n,i}\left(d(\theta_i^*||\theta_i) - d(\theta_i^*||\theta_i')\right)\right| \\
&\le \max_{1\le i\le k}|d(\theta_i^*||\theta_i) - d(\theta_i^*||\theta_i')| \\
&= \max_{1\le i\le k}|(\theta_i'-\theta_i)A'(\theta_i^*) + A(\theta_i) - A(\theta_i')| \\
&\le 2C\delta
\end{aligned}$$

where $C=\sup_{\theta\in(\underline{\theta},\overline{\theta})}|A'(\theta)|<\infty$, by Assumption 1. □

LEMMA EC.7. *For any open set* $\tilde{\Theta} \subset \Theta$,

$$\int\limits_{\boldsymbol{\theta}\in\tilde{\Theta}} \frac{\pi_n(\boldsymbol{\theta})}{\pi_n(\boldsymbol{\theta}^*)} d\boldsymbol{\theta} \doteq \int\limits_{\boldsymbol{\theta}\in\tilde{\Theta}} \exp\{-nW_n(\boldsymbol{\theta})\} d\boldsymbol{\theta}.$$

Proof. By Lemma EC.5, we can fix a sequence $\epsilon_n \geq 0$ with $\epsilon_n \to 0$ such that,

$$\exp\{-n(W_n(\boldsymbol{\theta})+\epsilon_n)\} \leq \frac{\pi_n(\boldsymbol{\theta})}{\pi_n(\boldsymbol{\theta}^*)} \leq \exp\{-n(W_n(\boldsymbol{\theta})-\epsilon_n)\}.$$

Integrating over $\tilde{\Theta}$ yields,

$$\exp\{-n\epsilon_n\} \int\limits_{\tilde{\Theta}} \exp\{-nW_n(\boldsymbol{\theta})\} d\boldsymbol{\theta} \leq \int\limits_{\tilde{\Theta}} \frac{\pi_n(\boldsymbol{\theta})}{\pi_n(\boldsymbol{\theta}^*)} d\boldsymbol{\theta} \leq \exp\{n\epsilon_n\} \int\limits_{\tilde{\Theta}} \exp\{-nW_n(\boldsymbol{\theta})\} d\boldsymbol{\theta}.$$

Taking the logarithm of each side implies

$$\frac{1}{n}\left|\log \int\limits_{\tilde{\Theta}} \frac{\pi_n(\boldsymbol{\theta})}{\pi_n(\boldsymbol{\theta}^*)} d\boldsymbol{\theta} - \log \int\limits_{\tilde{\Theta}} \exp\{-nW_n(\boldsymbol{\theta})\} d\boldsymbol{\theta}\right| \leq \epsilon_n \to 0.$$

□

LEMMA EC.8. *For any open set* $\tilde{\Theta} \subset \Theta$,

$$\int\limits_{\boldsymbol{\theta}\in\tilde{\Theta}} \exp\{-nW_n(\boldsymbol{\theta})\} d\boldsymbol{\theta} \doteq \exp\{-n \inf_{\boldsymbol{\theta}\in\tilde{\Theta}} W_n(\boldsymbol{\theta})\}$$

Proof. Let $\hat{\boldsymbol{\theta}}_n$ be a point in the closure of $\tilde{\Theta}$, satisfying

$$W_n(\hat{\boldsymbol{\theta}}_n) = \inf_{\boldsymbol{\theta}\in\tilde{\Theta}} W_n(\boldsymbol{\theta}).$$

Such a point always exists, since W_n is continuous, and the closure of $\tilde{\Theta}$ is compact. Let

$$\gamma_n \triangleq \int\limits_{\boldsymbol{\theta}\in\tilde{\Theta}} \exp\{-nW_n(\boldsymbol{\theta})\} d\boldsymbol{\theta}.$$

Our goal is to show

$$\frac{1}{n}\log(\gamma_n) + W_n(\hat{\boldsymbol{\theta}}_n) \to 0.$$

We have

$$\gamma_n \leq \mathrm{Vol}(\tilde{\Theta}) \exp\{-nW_n(\hat{\boldsymbol{\theta}}_n)\}$$

where for any $\Theta' \subset \Theta$, $\mathrm{Vol}(\Theta') = \int_{\tilde{\Theta}} d\boldsymbol{\theta} \in (0, \infty)$ denotes the volume of Θ. This shows

$$\limsup_{n\to\infty} \left(\frac{1}{n} \log(\gamma_n) + W_n(\hat{\boldsymbol{\theta}}_n) \right) \leq 0.$$

We now show the reverse. Fix an arbitrary $\epsilon > 0$. By Lemma EC.6, there exists $\delta > 0$ such that

$$|W_n(\boldsymbol{\theta}) - W_n(\hat{\boldsymbol{\theta}}_n)| \leq \epsilon \qquad \forall n \in \mathbb{N}$$

for any $\boldsymbol{\theta} \in \Theta$ with

$$\|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_n\|_\infty \leq \delta.$$

Now, choose a finite δ–cover O of $\tilde{\Theta}$ in the norm $\|\cdot\|_\infty$. Remove any set in O that does not intersect $\tilde{\Theta}$. Then, for each $o \in O$,

$$\mathrm{Vol}(o \cap \tilde{\Theta}) > 0 \implies C_\delta \triangleq \min_{o\in O} \mathrm{Vol}(o \cap \tilde{\Theta}) > 0.$$

Choose $o_n \in O$ with $\hat{\boldsymbol{\theta}}_n \in \mathrm{closure}(o_n)$. Then, for every $\boldsymbol{\theta} \in o_n$, $W_n(\boldsymbol{\theta}) \leq W_n(\hat{\boldsymbol{\theta}}_n) + \epsilon$. This shows

$$\gamma_n \geq \int_o \exp\{-nW_n(\boldsymbol{\theta})\} d\boldsymbol{\theta} \geq C_\delta \exp\{-n(W_n(\hat{\boldsymbol{\theta}}_n) - \epsilon)\}.$$

Taking the logarithm of both sides implies

$$\frac{1}{n} \log(\gamma_n) + W_n(\hat{\boldsymbol{\theta}}_n) \geq \frac{C_\delta}{n} - \epsilon \to -\epsilon.$$

Since ϵ was chosen arbitrarily, this shows

$$\liminf_{n\to\infty} \left(\frac{1}{n} \log(\gamma_n) + W_n(\hat{\boldsymbol{\theta}}_n) \right) \geq 0,$$

and completes the proof. $\square$

We now complete the proof of Proposition 5.

Proof of Proposition 5. We begin with a simple observation. For any sequences of real numbers $\{a_n\}, \{b_n\}$, and $\{\tilde{a}_n\}, \{\tilde{b}_n\}$, if $a_n \doteq \tilde{a}_n$ and $b_n \doteq \tilde{b}_n \in \mathbb{R}$, then $a_n/b_n \doteq \tilde{a}_n/\tilde{b}_n$.

Therefore, we have

$$\Pi_n(\tilde{\Theta}) = \frac{\Pi_n(\tilde{\Theta})}{\Pi_n(\Theta)} = \frac{\int_{\tilde{\Theta}} \pi_n(\boldsymbol{\theta}) d\boldsymbol{\theta}}{\int_{\Theta} \pi_n(\boldsymbol{\theta}) d\boldsymbol{\theta}} = \frac{\int_{\tilde{\Theta}} (\pi_n(\boldsymbol{\theta})/\pi_n(\boldsymbol{\theta}^*)) d\boldsymbol{\theta}}{\int_{\Theta} (\pi_n(\boldsymbol{\theta})/\pi_n(\boldsymbol{\theta}^*)) d\boldsymbol{\theta}} \doteq \frac{\exp\{-n \inf_{\boldsymbol{\theta}\in\tilde{\Theta}} W_n(\boldsymbol{\theta})\}}{\exp\{-n \inf_{\boldsymbol{\theta}\in\Theta} W_n(\boldsymbol{\theta})\}}$$

where the final equality follows from the previous two lemmas. Since $W_n(\boldsymbol{\theta}) \geq 0$ and $W_n(\boldsymbol{\theta}^*) = 0$, $\exp\{-n \inf_{\boldsymbol{\theta}\in\Theta} W_n(\boldsymbol{\theta})\} = 1$. $\square$

EC.6.4. Large Deviations of the Value Measure: Proof of Lemma 4

LEMMA 4. *For any $i \neq I^*$, $V_{n,i} \doteq \alpha_{n,i}$.*

The reader should glance back at Subsection 3.2 to recall the notation and assumptions regarding TTVS.

Proof. First, since

$$V_{n,i} = \int_{\Theta_i} v_i(\boldsymbol{\theta})\pi_n(\boldsymbol{\theta})d\boldsymbol{\theta} \leq (u(\overline{\theta}) - u(\underline{\theta})) \int_{\Theta_i} \pi_n(\boldsymbol{\theta})d\boldsymbol{\theta} = (u(\overline{\theta}) - u(\underline{\theta}))\alpha_{n,i}$$

it is immediate that

$$\limsup_{n\to\infty} n^{-1}(\log V_{n,i} - \log \alpha_{n,i}) \leq 0. \tag{EC.14}$$

The other direction is more subtle. Define $\Theta_{i,\delta} \subset \Theta_i$ by

$$\Theta_{i,\delta} = \{\boldsymbol{\theta} \in \Theta : \theta_i \geq \max_{j\neq i} \theta_j + \delta\}.$$

For any $\boldsymbol{\theta} \in \Theta_{i,\delta}$, $v_i(\boldsymbol{\theta}) \geq C_\delta$ where

$$C_\delta \equiv \min_{\theta\in[\underline{\theta},\overline{\theta}]} u(\theta + \delta) - u(\theta) > 0. \tag{EC.15}$$

Because $u(\theta + \delta) - u(\theta)$ is continuous and is strictly positive for each θ and $[\underline{\theta}, \overline{\theta}]$ is compact, this minimum exists and the objective value is strictly positive. Then

$$V_{n,i} \geq \int_{\Theta_{i,\delta}} v_i(\boldsymbol{\theta})\pi_n(\boldsymbol{\theta})d\boldsymbol{\theta} \geq C_\delta \int_{\Theta_{i,\delta}} \pi_n(\boldsymbol{\theta})d\boldsymbol{\theta} = C_\delta \Pi_n(\Theta_{i,\delta}) \qquad \forall \delta > 0.$$

Combining this with Proposition 5 shows

$$\liminf_{n\to\infty} \frac{1}{n}(\log V_{n,i} - \log \alpha_{n,i}) \geq \liminf_{n\to\infty} \frac{1}{n}(\log \Pi_n(\Theta_{i,\delta}) - \log \Pi_n(\Theta_i)) = -\min_{\boldsymbol{\theta}\in\Theta_{i,\delta}} D_{\overline{\psi}_n}(\boldsymbol{\theta}^*||\boldsymbol{\theta}) - \min_{\boldsymbol{\theta}\in\Theta_i} D_{\overline{\psi}_n}(\boldsymbol{\theta}^*||\boldsymbol{\theta}).$$

The final term can be made arbitrarily small by taking $\delta \to 0$. Precisely, by Lemma EC.6, for any $\epsilon > 0$, there exists $\delta > 0$ such that for all $n \in \mathbb{N}$ and $\boldsymbol{\theta}, \boldsymbol{\theta}' \in \Theta$ satisfying $\|\boldsymbol{\theta} - \boldsymbol{\theta}'\|_\infty \leq \delta$,

$$D_{\overline{\psi}_n}(\boldsymbol{\theta}^*||\boldsymbol{\theta}) \leq \epsilon.$$

Therefore, for each $\epsilon > 0$ one can choose $\delta > 0$ such that

$$\min_{\boldsymbol{\theta}\in\Theta_{i,\delta}} D_{\overline{\psi}_n}(\boldsymbol{\theta}^*||\boldsymbol{\theta}) \le \min_{\boldsymbol{\theta}\in\Theta_i} D_{\overline{\psi}_n}(\boldsymbol{\theta}^*||\boldsymbol{\theta}) + \epsilon.$$

This shows $\liminf n^{-1}(\log V_{n,i} - \log \alpha_{n,i}) \ge -\epsilon$ for all $\epsilon > 0$, and hence

$$\liminf_{n\to\infty} n^{-1}(\log V_{n,i} - \log \alpha_{n,i}) \ge 0.$$

□

EC.7. Simplifying and Bounding the Error Exponent

EC.7.1. Proof of Lemma 2

To begin, we restate the results of Lemma 2 in the order in which they will be proved. Recall, from Section EC.5 that $A(\theta)$ is increasing and strictly convex, and, by (EC.3), $A'(\theta)$ is the mean observation under θ.

LEMMA 2 *Define for each $i \ne I^*, \psi \ge 0$,*

$$C_i(\beta,\psi) \triangleq \min_{x\in\mathbb{R}} \beta d(\theta^*_{I^*}||x) + \psi d(\theta^*_i||x). \tag{EC.16}$$

(a) *For any $i \ne I^*$ and probability distribution $\boldsymbol{\psi}$ over $\{1,\dots,k\}$*

$$\min_{\boldsymbol{\theta}\in\overline{\Theta}_i} D_{\boldsymbol{\psi}}(\boldsymbol{\theta}^*||\boldsymbol{\theta}) = C_i(\psi_{I^*},\psi_i).$$

where $\overline{\Theta}_i \triangleq \{\boldsymbol{\theta}\in\Theta | \theta_i \ge \theta_{I^*}\}$.

(b) *Each C_i is a concave function.*

(c) *The unique solution to the minimization problem (EC.16) is $\overline{\theta}\in\mathbb{R}$ satisfying*

$$A'(\overline{\theta}) = \frac{\psi_{I^*}A'(\theta^*_{I^*}) + \psi_i A'(\theta^*_i)}{\psi_{I^*}+\psi_i}.$$

Therefore,

$$C_i(\psi_{I^*},\psi_i) = \psi_{I^*} d(\theta^*_{I^*}||\overline{\theta}) + \psi_i d(\theta^*_i||\overline{\theta}).$$

(d) *Each C_i is a strictly increasing function.*

Proof. **(a)**

$$
\begin{aligned}
\min_{\boldsymbol{\theta}\in\overline{\Theta}_i} D_{\boldsymbol{\psi}}(\boldsymbol{\theta}^*||\boldsymbol{\theta}) &= \min_{\boldsymbol{\theta}\in\Theta:\theta_i\geq\theta_{I^*}} \sum_{j=1}^{k} \psi_{n,j} d(\theta^*_j||\theta_j) \\
&= \min_{\overline{\theta}\geq\theta_i\geq\theta_{I^*}\geq\underline{\theta}} \psi_{I^*} d(\theta^*_{I^*}||\theta_{I^*}) + \psi_i d(\theta^*_i||\theta_i) + \sum_{j\notin\{i,I^*\}} \min_{\theta_j} \psi_{n,j} d(\theta^*_j||\theta_j) \\
&= \min_{\overline{\theta}\geq\theta_i\geq\theta_{I^*}\geq\underline{\theta}} \psi_{I^*} d(\theta^*_{I^*}||\theta_{I^*}) + \psi_i d(\theta^*_i||\theta_i)
\end{aligned}
$$

where the last equality uses that the minimum occurs when $\theta_j = \theta^*_j$ for $j \notin \{I^*, i\}$, and this is feasible for any choice of (θ_i, θ_{I^*}). Then, by the monotonicity properties of KL-divergence (see Section EC.5, equation (EC.5)), there is always a minimum with $\theta_i = \theta_{I^*}$. Therefore this objective value is equal to

$$
\min_{\theta\in[\underline{\theta},\overline{\theta}]} \psi_{I^*} d(\theta^*_{I^*}||\theta) + \psi_i d(\theta^*_i||\theta) = \min_{x\in\mathbb{R}} \psi_{I^*} d(\theta^*_{I^*}||x) + \psi_i d(\theta^*_i||x) = C_i(\psi_{I^*}, \psi_i).
$$

(b) C_i is the minimum over a family of linear functions and therefore is concave (See Chapter 3.2 of Boyd and Vandenberghe (2004)). In particular $C_i(\beta,\psi) = \min_{x\in\mathbb{R}} g((\beta,\psi);x)$ where $g((\beta,\psi);x) = \beta d(\theta^*_{I^*}||x) + \psi d(\theta^*_i||x)$ is linear in (β,ψ).

(c) Direct calculation using the formula for KL divergence in exponential families (see (EC.4) in Section EC.5) shows

$$
\beta d(\theta^*_{I^*}||x) + \psi_i d(\theta^*_i||x) = (\beta+\psi_i)A(x) - (\beta A'(\theta^*_{I^*}) + \psi_i A'(\theta^*_i))x + f(\beta, \theta^*_{I^*}, \psi_i, \theta^*_i)
$$

where $f(\beta, \theta^*_{I^*}, \psi_i, \theta^*_i)$ captures terms that are independent of x. Setting the derivative with respect to x to zero yields the result since $A(x)$ is strictly convex.

(d) We will show C_i is strictly increasing in the second argument. The proof that it is strictly increasing in its first argument follows by symmetry. Set

$$
f(\psi_i, x) = \beta d(\theta^*_{I^*}||x) + \psi_i d(\theta^*_i||x)
$$

so that $C_i(\beta,\psi_i)=\min_{x\in\mathbb{R}} f(\psi_i,x)$. Since KL divergences are non-negative, $f(\psi_i,x)$ is weakly increasing in ψ_i. To establish the claim, fix two nonnegative numbers $\psi'<\psi''$. Let $x'=\arg\min_x f(\psi',x)$ and $x''=\arg\min_x f(\psi'',x)$. By part (c), these are unique and $x'<x''$. Then

$$C_i(\beta,\psi')=f(\psi',x')<f(\psi',x'')\le f(\psi'',x'')=C_i(\beta,\psi'')$$

where the first inequality uses that $x'\neq x''$ and x' is a unique minimum and the second uses the f is non-decreasing. □

EC.7.2. Proof of Proposition 7

We will begin by restating Proposition 7.

PROPOSITION 7. *The solution to the optimization problem* (12) *is the unique allocation* $\boldsymbol{\psi}^*$ *satisfying* $\psi^*_{I^*}=\beta$ *and*

$$C_i(\beta,\psi_i^*)=C_j(\beta,\psi_j^*)\qquad \forall i,j\neq I^*. \tag{EC.17}$$

If $\boldsymbol{\psi}_n=\boldsymbol{\psi}^*$ *for all* n*, then*

$$\Pi_n(\Theta^c_{I^*})\doteq\exp\{-n\Gamma^*_\beta\}.$$

Moreover under any other adaptive allocation rule, if $\overline{\psi}_{n,I^*}\to\beta$ *as* $n\to\infty$ *then*

$$\limsup_{n\to\infty}-\frac{1}{n}\log\Pi_n(\Theta^c_{I^*})\le\Gamma^*_\beta$$

almost surely.

Proof. By Lemma 2, each function C_i is continuous, and therefore $\min_{i\neq I^*}C_i(\beta,\psi_i)$ is continuous in $(\psi_i:i\neq I^*)$. Since continuous functions on a compact space attain their minimum, there exists an optimal solution $\boldsymbol{\psi}^*$ to (12), which by definition satisfies

$$\min_{i\neq I^*}C_i(\beta,\psi_i^*)=\max_{\boldsymbol{\psi}:\psi_{I^*}=\beta}\min_{i\neq I^*}C_i(\beta,\psi_i).$$

Suppose $\boldsymbol{\psi}^*$ does not satisfy (EC.17), so for some $j\neq I^*$,

$$C_j(\beta,\psi_j^*)>\min_{i\neq I^*}C_i(\beta,\psi_i^*).$$

This yields a contradiction. Consider a new vector $\boldsymbol{\psi}^{\epsilon}$ with $\boldsymbol{\psi}_j^{\epsilon} = \psi_j^* - \epsilon$ and $\psi_i^{\epsilon} = \psi_i^* + \epsilon/(k-2)$ for each $i \notin \{I^*, j\}$. For sufficiently small ϵ, one has

$$C_j(\beta, \psi_j^{\epsilon}) > \min_{i \neq I^*} C_i(\beta, \psi_i^{\epsilon}) > \min_{i \neq I^*} C_i(\beta, \psi_i^*)$$

and so $\boldsymbol{\psi}^{\epsilon}$ attains a higher objective value. To show the solution to (EC.17) must be unique, imagine $\boldsymbol{\psi}$ and $\boldsymbol{\psi}'$ both satisfy (EC.17) and $\psi_{I^*} = \psi'_{I^*} = \beta$. If $\psi_j > \psi'_j$ for some j, then $C_j(\beta, \psi_j) > C_j(\beta, \psi'_j)$ since C_j is strictly increasing. But by (EC.17) this implies that $C_j(\beta, \psi_j) > C_j(\beta, \psi'_j)$ for *every* $j \neq I^*$, which implies $\psi_j > \psi'_j$ for every j, and contradicts that that $\sum_{j \neq I^*} \psi_j = \sum_{j \neq I^*} \psi'_j = 1 - \beta$.

The remaining claims follow immediately from Propoosition 5 and Lemma 2, which together show that under any adaptive allocation rule

$$\Pi_n(\Theta_{I^*}^c) \doteq \exp\{-n \min_{i \neq I^*} C_i(\overline{\psi}_{n,I^*}, \overline{\psi}_{n,i})\}.$$

This implies that if $\overline{\boldsymbol{\psi}}_n = \boldsymbol{\psi}^*$ for all n, then $\Pi_n(\Theta_{I^*}^c) \doteq \exp\{-n\Gamma_\beta^*\}$. Similarly, by the continuity of each C_i, if $\overline{\psi}_{n,I^*} \to \beta$, then

$$\Pi_n(\Theta_{I^*}^c) \doteq \exp\{-n \min_{i \neq I^*} C_i(\beta, \overline{\psi}_{n,i})\} \geq \exp\{-n\Gamma_\beta^*\}$$

which establishes the final claim. □

EC.7.3. Proof of Lemma 3

Recall, the notation

$$\Gamma^* = \max_{\boldsymbol{\psi}} \min_{i \neq I^*} C_i(\psi_{I^*}, \psi_i) \qquad \Gamma_\beta^* \triangleq \max_{\boldsymbol{\psi}: \psi_{I^*} = \beta} \min_{i \neq I^*} C_i(\beta, \psi_i)$$

where

$$C_i(\beta, \psi) = \min_{x \in \mathbb{R}} \beta d(\theta_{I^*}^* || x) + \psi d(\theta_{i^*}^* || x).$$

LEMMA 3. *For* $\beta^* = \arg\max_\beta \Gamma_\beta^*$ *and any* $\beta \in (0,1)$,

$$\frac{\Gamma^*}{\Gamma_\beta^*} \leq \max\left\{\frac{\beta^*}{\beta}, \frac{1-\beta^*}{1-\beta}\right\}.$$

Therefore $\Gamma^* \leq 2\Gamma_{1/2}^*$.

Proof. Define for each non-negative vector $\boldsymbol{\psi}$,

$$f(\boldsymbol{\psi}) = \min_{i \neq I^*} C_i(\psi_{I^*}, \psi_i).$$

The optimal exponent Γ^* is the maximum of $f(\boldsymbol{\psi})$ over probability vectors $\boldsymbol{\psi}$. Here, we instead define f for all non-negative vectors, and proceed by varying the total budget of measurement effort available $\sum_{i=1}^k \psi_i$.

Because each C_i is non-decreasing (see Lemma 2), f is non-decreasing. Since the minimum over x in the definition of C_i only depends on the relative size of the components of $\boldsymbol{\psi}$, f is homogenous of degree 1. That is $f(c\boldsymbol{\psi}) = cf(\boldsymbol{\psi})$ for all $c \geq 1$. For each $c_1, c_2 > 0$ define

$$g(c_1, c_2) = \max\{f(\boldsymbol{\psi}) : \psi_{I^*} = c_1, \sum_{i \neq I^*} \psi_i \leq c_2, \boldsymbol{\psi} \geq 0\}.$$

The function g inherits key properties of f; it is also non-decreasing and homogenous of degree 1. We have

$$\begin{aligned}
\Gamma^*_\beta &= \max\{f(\boldsymbol{\psi}) : \psi_{I^*} = \beta, \sum_{i=1}^k \psi_i = 1, \boldsymbol{\psi} \geq 0\} \\
&= \max\{f(\boldsymbol{\psi}) : \psi_{I^*} = \beta, \sum_{i \neq I^*} \psi_i \leq 1 - \beta, \boldsymbol{\psi} \geq 0\} \\
&= g(\beta, 1 - \beta)
\end{aligned}$$

where the second equality uses that f is non-decreasing. Similarly, $\Gamma^* = g(\beta^*, 1 - \beta^*)$. Setting

$$r := \max\left\{\frac{\beta^*}{\beta}, \frac{1 - \beta^*}{1 - \beta}\right\}$$

implies $r\beta \geq \beta^*$ and $r(1 - \beta) \geq 1 - \beta^*$. Therefore

$$r\Gamma^*_\beta = rg(\beta, 1 - \beta) = g(r\beta, r(1 - \beta)) \geq g(\beta^*, 1 - \beta^*) = \Gamma^*.$$

□

EC.7.4. Sub-Gaussian Bound: Proof of Proposition 1

The proof of Proposition 1 relies on the following variational form of Kullback–Leibler divergence, which is given in Theorem 5.2.1 of Robert Gray's textbook *Entropy and Information Theory* Gray (2011).

LEMMA EC.9. *Fix two probability measures* $\mathbf{P}$ *and* $\mathbf{Q}$ *defined on a common measureable space* $(\Omega, \mathcal{F})$. *Suppose that* $\mathbf{P}$ *is absolutely continuous with respect to* $\mathbf{Q}$. *Then*

$$D\left(\mathbf{P}||\mathbf{Q}\right) = \sup_{X} \left\{\mathbb{E}_{\mathbf{P}}[X] - \log \mathbb{E}_{\mathbf{Q}}[e^{X}]\right\},$$

where the supremum is taken over all random variables X *such that the expectation of* X *under* $\mathbf{P}$ *is well defined, and* e^X *is integrable under* $\mathbf{Q}$.

When comparing two normal distributions $\mathcal{N}(\theta, \sigma^2)$ and $\mathcal{N}(\theta', \sigma^2)$ with common variance, the KL-divergence can be expressed as $d(\theta||\theta') = (\theta - \theta')^2/(2\sigma^2)$. We follow Russo and Zou (2015) in deriving the following corollary of Fact EC.9, which provides and analogous lower bound on the KL-divergences when distributions are sub-Gaussian. Recall that, $\mu(\theta) = \int y p(y|\theta) d\nu(y)$ denotes the mean observation under θ.

COROLLARY EC.3. *Fix any* $\theta, \theta' \in [\underline{\theta}, \overline{\theta}]$. *If when* $Y \sim p(y|\theta')$, Y *is sub-Gaussian with parameter* σ, *then,*

$$d(\theta||\theta') \geq \frac{\left(\mu(\theta) - \mu(\theta')\right)^2}{2\sigma^2}.$$

Proof. Consider two alternate probability distributions for a random variable Y, one where $Y \sim p(y|\theta)$ and one where $Y \sim p(y|\theta')$. We apply Fact $EC.9$ where $X = \lambda(Y - \mathbb{E}_{\theta'}[Y])$, $\mathbf{P}$ is the probability measure when $Y \sim p(y|\theta)$ and $\mathbf{Q}$ is the measure when $Y \sim p(y|\theta')$. By the sub-Gaussian assumption $\log \mathbb{E}_{\theta'}\left[\exp\{X\}\right] \leq \lambda^2\sigma^2/2$. Therefore, Fact $EC.9$ implies

$$d(\theta||\theta') \geq \lambda(\mathbb{E}_{\theta}[X]) - \frac{\lambda^2\sigma^2}{2} = \lambda(\mathbb{E}_{\theta}[Y] - \mathbb{E}_{\theta'}[Y]) - \frac{\lambda^2\sigma^2}{2}.$$

The result follows by choosing $\lambda = (\mathbb{E}_{\theta}[Y] - \mathbb{E}_{\theta'}[Y])/\sigma^2$ which minimizes the right hand side. $\square$

We are now ready to prove Proposition 1. Recall that in an exponential family, $A'(\theta) = \int T(y)p(y|\theta)d\nu(y)$, so if $T(y) = y$ then $A'(\theta) = \mu(\theta)$.

Proof of Proposition 1. By Lemma 2,

$$\Gamma^*_{1/2} = \max_{\boldsymbol{\psi}:\psi_{I^*}=1/2} \min_{i\neq I^*} C_i(1/2, \psi_i)$$

Let $\mu_{I^*} = A'(\theta^*_{I^*})$ and $\mu_i = A'(\theta^*_i)$ denote the means of designs I^* and i so $\Delta_i = \mu_{I^*} - \mu_i$. By Lemma 2,

$$C_i(1/2, \psi_i) = (1/2)d(\theta^*_{I^*}||\overline{\theta}) + \psi_i d(\theta^*_i||\overline{\theta}).$$

where $\overline{\theta}$ is the unique parameter with mean

$$A'(\overline{\theta}) = \frac{(1/2)\mu_{I^*} + \psi_i\mu_i}{1/2 + \psi_i}.$$

For $\psi_i \leq 1/2$,

$$A'(\overline{\theta}) \geq \frac{\mu_{I^*} + \mu_i}{2} = \mu_i + \Delta_i/2.$$

Now, using Corollary EC.3 and the non-negativity of KL-divergence

$$C_i(1/2, \psi_i) \geq \psi_i d(\theta^*_i||\overline{\theta}) \geq \frac{\psi_i\left(\mu_i - (\mu_i - \Delta_i/2)\right)^2}{2\sigma^2} = \frac{\psi_i \Delta_i^2}{8\sigma^2}.$$

To lower bound $\Gamma^*_{1/2}$ it is sufficient to construct some $\boldsymbol{\psi}$ with high objective value. Choosing $\psi_{I^*} = 1/2$, and $\psi_i \propto \Delta_i^{-2}$, so

$$\psi_i = \frac{1}{2}\left(\sum_{j\neq I^*} \Delta_j^{-2}\right)^{-1} \Delta_i^{-2}$$

yields

$$\min_{i\neq I^*} C_i(1/2, \psi_i) \geq \frac{1}{16\sigma^2 \sum_2^k \Delta_j^{-2}}.$$

□

EC.7.5. Convergence of Uniform Allocation: Proof of Proposition 2

Proof. Without loss of generality, assume the problem is parameterized so that the mean of design i is θ_i^* By Proposition 6, we have

$$\Pi_n(\Theta_{I^*}^c) \doteq \exp\{-n \min_{i \neq I^*} C_i(k^{-1}, k^{-1})\}$$

By Lemma 2,

$$C_i(k^{-1}, k^{-1}) = k^{-1} d(\theta_{I^*}^* || \overline{\theta}) + k^{-1} d(\theta_i^* || \overline{\theta})$$

where $\overline{\theta} = (\theta_{I^*}^* + \theta_i^*)/2$. Therefore, using the formula for the KL-divergence of standard Gaussian random variables,

$$C_i(k^{-1}, k^{-1}) = \frac{(\theta_{I^*}^* - \overline{\theta})^2}{2\sigma^2} + \frac{(\theta_i^* - \overline{\theta})^2}{2\sigma^2} = \frac{(\theta_{I^*}^* - \theta_i^*)^2}{4\sigma^2} = \frac{\Delta_i^2}{4\sigma^2}.$$

□

EC.8. Analysis of the Top-Two Allocation Rules: Proof of Proposition 8

PROPOSITION 8. *Under the TTTS, TTPS, or TTVS algorithm with parameter $\beta > 0$, $\overline{\boldsymbol{\psi}}_n \to \boldsymbol{\psi}^\beta$, where $\boldsymbol{\psi}^\beta$ is the unique allocation with $\psi_{I^*}^\beta = \beta$ satisfying*

$$C_i(\beta, \psi_i^\beta) = C_j(\beta, \psi_j^\beta) \qquad \forall i, j \neq I^*. \tag{EC.18}$$

Therefore,

$$\Pi_n(\Theta_{I^*}^c) \doteq e^{-n\Gamma_\beta^*}. \tag{EC.19}$$

Because each C_i is continuous, if $\overline{\boldsymbol{\psi}}_n \to \boldsymbol{\psi}^\beta$ then $C_i(\overline{\psi}_{n,I^*}, \overline{\psi}_{n,i}) \to C_i(\beta, \psi_i^\beta)$ for all $i \neq I^*$. Equation (EC.19) then follows by invoking Proposition 7, which establishes the optimality of the allocation $\boldsymbol{\psi}^\beta$.

The remainder of this section establishes that $\overline{\boldsymbol{\psi}}_n \to \boldsymbol{\psi}^\beta$ almost surely under the proposed top-two rules. The proof is broken into a number of steps. In order to provide a nearly unified treatment of the three algorithms, we begin with several results that hold for any allocation rule.

EC.8.1. Results for a general allocation rule

As in other sections, all arguments here hold for any sample path (up to a set of measure zero). The first result provides a sufficient condition under which $\overline{\boldsymbol{\psi}}_n \to \boldsymbol{\psi}^\beta$. Roughly speaking, if $\overline{\psi}_{n,j} \geq \psi_j^\beta + \delta$, then too much measurement effort has been allocated to design j relative to the optimal proportion ψ_j^β. Algorithms satisfying (EC.20) allocate negligible measurement effort to such designs, and therefore the average measurement effort they receive must decrease toward the optimal proportion.

Lemma EC.10 **(Sufficient condition for optimality)**. *Consider any adaptive allocation rule. If* $\overline{\psi}_{n,I^*} \to \beta$ *and*

$$\sum_{n\in\mathbb{N}} \psi_{n,j} \mathbf{1}(\overline{\psi}_{n,j} \geq \psi_j^\beta + \delta) < \infty \qquad \forall\, j \neq I^*,\, \delta > 0, \tag{EC.20}$$

then $\overline{\boldsymbol{\psi}}_n \to \boldsymbol{\psi}^\beta$.

Proof. Fix a sample path for which $\psi_{n,I^*} \to \beta$, and (EC.20) holds. Fix some $j \neq I^*$. We first show $\liminf_{n\to\infty} \overline{\psi}_{n,j} \leq \psi_j^*$. Suppose otherwise. Then, with positive probability, for some $\delta > 0$, there exists N such that for all $n \geq N$, $\overline{\psi}_{n,j} \geq \psi_j^* + \delta$. But then,

$$\sum_{n\in\mathbb{N}} \psi_{n,j} = \sum_{n=1}^{N} \psi_{n,j} + \sum_{n=N+1}^{\infty} \mathbf{1}(\overline{\psi}_{n,j} \geq \psi_j^* + \delta)\psi_{n,j} < \infty.$$

But since $\overline{\psi}_{n,j} = \sum_{\ell=1}^n \psi_{n,j}/n$ this implies $\overline{\psi}_{n,j} \to 0$.

Now, we show $\limsup_{n\to\infty} \overline{\psi}_{n,j} \leq \psi_j^*$. Proceeding by contradiction again, suppose otherwise. Then, with positive probability

$$\limsup_{n\to\infty} \overline{\psi}_{n,j} > \psi_j^\beta \qquad \& \qquad \liminf_{n\to\infty} \overline{\psi}_{n,j} \leq \psi_j^\beta.$$

On any sample path where this occurs, for some $\delta > 0$, there exists an infinite sequence of times $N_1 < N_2 < N_3 < \ldots$ such that $\overline{\psi}_{N_\ell,j} \geq \psi_j^\beta + 2\delta$ when ℓ is odd and $\overline{\psi}_{N_\ell,j} \leq \psi_j^\beta + \delta$ when ℓ is even. This can only occur if,

$$\sum_{n\in\mathbb{N}} \psi_{n,j} \mathbf{1}(\overline{\psi}_{n,j} \geq \psi_j^* + \delta) = \infty,$$

which violates the hypothesis.

Together with the hypothesis that $\overline{\psi}_{n,I^*} \to \beta$, this implies that for all $i \in \{1, \ldots, k\}$, $\limsup_{n\to\infty} \overline{\psi}_{n,i} \leq \psi_i^\beta$. But since $\sum_i \overline{\psi}_{n,i} = \sum_i \psi_i^\beta$, this implies $\overline{\boldsymbol{\psi}}_n \to \boldsymbol{\psi}^\beta$. $\square$

The next lemma will be used to establish that (EC.20) holds for each of the proposed algorithms. It shows that if too much measurement effort has been allocated to some design $i \neq I^*$, in the sense that $\overline{\psi}_{n,i} > \psi_i^\beta + \delta$ for a constant $\delta > 0$, then $\alpha_{n,i}$ is exponentially small compared $\max_{j \neq I^*} \alpha_{n,j}$.

LEMMA EC.11 **(Over-allocation implies negligible probability)**. *Fix any $\delta > 0$ and $j \neq I^*$. With probability 1, under any allocation rule, if $\overline{\psi}_{n,I^*} \to \beta$, there exists $\delta' > 0$ and a sequence ϵ_n with $\epsilon_n \to 0$ such that for any $n \in \mathbb{N}$,*

$$\overline{\psi}_{n,j} \geq \psi_j^\beta + \delta \implies \frac{\alpha_{n,j}}{\max_{i \neq I^*} \alpha_{n,i}} \leq e^{-n(\delta' + \epsilon_n)}.$$

Proof. Since $\Pi_n(\Theta_{I^*}^c) = \sum_{i \neq I^*} \alpha_{n,i}$, $\Pi_n(\Theta_{I^*}^c) \doteq \max_{i \neq I^*} \alpha_{n,i}$. Then, by invoking Proposition (7), since $\overline{\boldsymbol{\psi}}_{n,I^*} \to \beta$,

$$\limsup_{n \to \infty} \; -\frac{1}{n} \log \left(\max_{i \neq I^*} \alpha_{n,i} \right) \leq \Gamma_\beta^*.$$

Recall the definition $\overline{\Theta}_i \triangleq \{\boldsymbol{\theta} | \theta_i \geq \theta_{I^*}\}$. Now, by Proposition 5 and Lemma 2,

$$\alpha_{n,j} = \Pi_n(\Theta_j) \leq \Pi_n(\overline{\Theta}_j) \doteq \exp\{-nC_j(\overline{\psi}_{n,I^*}, \overline{\psi}_{n,j})\} \doteq \exp\{-nC_j(\beta, \overline{\psi}_{n,j})\}.$$

Combining these equations implies that there exists a non-negative sequence $\epsilon_n \to 0$ with

$$\frac{\alpha_{n,j}}{\max_{i \neq I^*} \alpha_{n,i}} \leq \frac{\exp\{-n(C_j(\beta, \overline{\psi}_{n,j}) - \epsilon_n/2)\}}{\exp\{-n(\Gamma_\beta^* + \epsilon_n/2)\}} = \exp\left\{-n\left((C_j(\beta, \overline{\psi}_{n,j}) - \Gamma_\beta^*) - \epsilon_n\right)\right\}.$$

Since $C_j(\beta, \psi_j)$ is strictly increasing in ψ_j (See lemma 2) and $C_j(\beta, \psi_j^\beta) = \Gamma_\beta^*$, there exists some $\delta' > 0$ such that

$$\overline{\psi}_{n,j} \geq \psi_j^\beta + \delta \implies C_j(\beta, \overline{\psi}_{n,j}) - \Gamma_\beta^* > \delta'.$$

□

The next result builds on Proposition 4. It shows that the quality of any design which receives infinite measurement effort is identified to arbitrary precision. On the other hand, for designs receiving finite measurement effort, there is always nonzero probability under the posterior that one of them significantly exceeds the highest quality that has been confidently identified. Therefore, $\alpha_{n,i}$ and $V_{n,i}$ remain bounded away from 0 for designs that receive finite measurement effort. This result will be used to show that all designs receive infinite measurement effort under the proposed top-two allocation rules, and as a result the posterior converges on the truth asymptotically.

Lemma EC.12 **(Implications of finite measurement)**. *Let*

$$\mathcal{I} = \{i \in \{1,..,k\} : \sum_{n=1}^{\infty} \psi_{n,i} < \infty\}$$

denote the set of designs to which a finite amount of measurement effort is allocated. Then, for any $i \notin \mathcal{I}$

$$\Pi_n\left(\{\boldsymbol{\theta} : \theta_i \in (\theta_i^* - \epsilon, \theta_i^* + \epsilon)\right) \to 1, \tag{EC.21}$$

and if $\mathcal{I}$ *is empty*

$$V_{n,i} \to \begin{cases} 0 & \textit{if } i \neq I^* \\ v_{I^*}(\boldsymbol{\theta}^*) > 0 & \textit{if } i = I^* \end{cases} \quad \text{and} \quad \alpha_{n,i} \to \begin{cases} 0 & \textit{if } i \neq I^* \\ 1 & \textit{if } i = I^*. \end{cases} \tag{EC.22}$$

If $\mathcal{I}$ *is nonempty, then for every* $i \in \mathcal{I}$,

$$\liminf_{n\to\infty} \alpha_{n,i} > 0 \quad \text{and} \quad \liminf_{n\to\infty} V_{n,i} > 0. \tag{EC.23}$$

Proof. Equation (EC.21) is implied by Proposition 4. Equation (**??**) is an almost immediate consequence of the convergence of Π_n to a point-mass at $\boldsymbol{\theta}^*$, again shown in Prop 4. We turn to show (EC.23), which is more subtle. Now, set

$$\Theta_{i,\epsilon} = \{\boldsymbol{\theta} \in \Theta : \theta_i \geq \max_{j \neq i} \theta_j + \epsilon\}$$

to be the set of parameters under which the quality of design i exceeds that of all others by at least ϵ. Let $\rho^* = \max_{i \notin \mathcal{I}} \theta_i^*$ denote the quality of the best design among those that are sampled infinitely often, and choose $\epsilon > 0$ small enough that $\rho^* + 2\epsilon < \overline{\theta}$. For $i \in \mathcal{I}$, we have

$$\Pi_n(\Theta_{i,\epsilon}) \geq \Pi_n(A) - \Pi_n(B)$$

for

$$A \equiv \{\boldsymbol{\theta} | \theta_i \geq \rho^* + 2\epsilon \;\&\; \theta_j < \rho^* \; \forall j \in \mathcal{I} \setminus \{i\}\}$$

defined to be parameters under which $\theta_i \geq \rho^* + 2\epsilon$ but none of the other designs in $\mathcal{I}$ exceed ρ^*, and

$$B \equiv \{\boldsymbol{\theta} : \max_{i \notin \mathcal{I}} \theta_i \geq \rho^* + \epsilon\}$$

defined to be the parameter vectors under which there is no design in $\mathcal{I}^c$ with quality exceeding $\rho^* + \epsilon$. By (EC.21),

$$\Pi_n(B) \to 0,$$

but by the second part of Proposition 4, the set of parameters A cannot be completely ruled based on a finite amount of measurement effort, and

$$\inf_{n\in\mathbb{N}} \Pi_n(A) > 0.$$

Together this shows

$$\liminf_{n\to\infty} \Pi_n(\Theta_{i,\epsilon}) > 0.$$

This implies the result, since $\alpha_{n,i} \geq \Pi_n(\Theta_{i,\epsilon})$ and $V_{n,i} \geq C_\epsilon \Pi_n(\Theta_{i,\epsilon})$ where as in (EC.15) we define

$$C_\epsilon \equiv \min_{\theta\in[\underline{\theta},\overline{\theta}-\epsilon]} v(\theta+\epsilon) - v(\theta) > 0.$$

$\square$

EC.8.2. Results specific to the proposed algorithms

We now leverage the general results of the previous subsection to show $\overline{\boldsymbol{\psi}} \to \boldsymbol{\psi}^\beta$ under each proposed top-two allocation rule. Proofs are provided separately for each of the three algorithms, but they follow a similar structure. In the first step, we use Lemma EC.12 to argue that $\overline{\psi}_{n,I^*} \to \beta$ almost surely. The proof then uses Lemma EC.11 to show (EC.20) holds, which by Lemma EC.10 is sufficient to establish that $\overline{\boldsymbol{\psi}}_n \to \boldsymbol{\psi}^\beta$.

EC.8.2.1. Top-Two Thompson Sampling Recall from Subsection 3.4, that under top-two Thompson sampling, for every $i \in \{1,...,k\}$,

$$\psi_{n,i} = \alpha_{n,i}\left(\beta + (1-\beta)\sum_{j\neq i}\frac{\alpha_{n,j}}{1-\alpha_{n,j}}\right).$$

Proof for TTTS. *Step 1: Show $\overline{\psi}_{n,I^*} \to \beta$.* To begin, we show $\sum_{n\in\mathbb{N}} \psi_{n,i} = \infty$ for each design i. Suppose otherwise. Let $\mathcal{I} = \{i \in \{1,..,k\} : \sum_1^\infty \psi_{n,i} < \infty\}$ be the set of designs to which finite measurement effort is allocated. Under the TTTS sampling rule, $\psi_{n,i} \geq \beta\alpha_{n,i}$. By Lemma EC.12, if $i \in \mathcal{I}$ then $\liminf_{n\to\infty} \alpha_{n,i} > 0$, which therefore implies $\sum_{n\in\mathbb{N}} \psi_{n,i} = \infty$, a contradiction.

Since $\sum_1^\infty \psi_{n,i} = \infty$ for all i, by applying Lemma EC.12 we conclude that $\alpha_{n,I^*} \to 1$. For TTTS, this implies $\psi_{n,I^*} \to \beta$ and therefore $\overline{\psi}_{n,I^*} \to \beta$.

Step 2: Show (EC.20) *holds.* By Lemma EC.10, it is enough to show that (EC.20) holds under TTTS. Let $\hat{I}_n = \arg\max_i \alpha_{n,i}$, and $\hat{J}_n = \arg\max_{i\neq \hat{I}_n} \alpha_{n,i}$. Since $\alpha_{n,I^*} \to 1$, for each sample path there is a finite time $\tau < \infty$ such that for all $n \geq \tau$, $\hat{I}_n = I^*$ and therefore $\hat{J}_n = \arg\max_{i\neq I^*} \alpha_{n,i}$. Under TTTS,

$$\psi_{n,i} \leq \alpha_{n,i}\left(\beta + \frac{1-\beta}{\alpha_{n,J_n}}\right) \leq \frac{\alpha_{n,i}}{\alpha_{n,J_n}},$$

where the first inequality follows since

$$\sum_{j\neq i} \frac{\alpha_{n,j}}{1-\alpha_{n,j}} \leq \frac{\sum_{j\neq i} \alpha_{n,i}}{1-\alpha_{n,\hat{I}_n}} \leq \frac{\sum_{j\neq i} \alpha_{n,j}}{\alpha_{n,\hat{J}_n}} \leq \frac{1}{\alpha_{n,\hat{J}_n}}.$$

For $n \geq \tau$, this means $\psi_{n,i} \leq \alpha_{n,i}/(\max_{j\neq I^*} \alpha_{n,j})$ for any $i \neq I^*$. By Lemma EC.11, there is a constant $\delta' > 0$ and a sequence $\epsilon_n \to 0$ such that

$$\overline{\psi}_{n,i} \geq \psi_i^\beta + \delta \implies \frac{\alpha_{n,i}}{\max_{j\neq I^*} \alpha_{n,j}} \leq e^{-n(\delta'-\epsilon_n)}.$$

Therefore for all $i \neq I^*$

$$\sum_{n\geq\tau} \psi_{n,i}\mathbf{1}(\overline{\psi}_{n,i} \geq \psi_i^\beta + \delta) \leq \sum_{n\geq\tau} e^{-n(\delta'-\epsilon_n)} < \infty.$$

□

Recall that top-two probability sampling sets $\psi_{n,\hat{I}_n} = \beta$ and $\psi_{n,\hat{J}_n} = 1-\beta$ where $\hat{I}_n = \arg\max_i \alpha_{n,i}$ and $\hat{J}_n = \arg\max_{j\neq \hat{I}_n} \alpha_{n,i}$ are the two designs with the highest posterior probability of being optimal.

Proof for TTPS. *Step 1: Show* $\overline{\psi}_{n,I^*} \to \beta$. To begin, we show $\sum_{n\in\mathbb{N}} \psi_{n,i} = \infty$ for each design i. Suppose otherwise. Let $\mathcal{I} = \{i \in \{1,..,k\} : \sum_1^\infty \psi_{n,i} < \infty\}$ be the set of designs to which finite measurement effort is allocated. Proceeding by contradiction, suppose $\mathcal{I}$ is nonempty. By Lemma

EC.12, there is a time τ and some probability $\alpha' > 0$ such that $\alpha_{n,i} > \alpha'$ for all $n \geq \tau$ and $i \in \mathcal{I}$. However, because of the assumption that $\theta_i^* \neq \theta_j^*$, for $i \neq j$, $I = \arg\max_{i \notin \mathcal{I}} \theta_i^*$ is unique. By (EC.21), the algorithm identifies $\arg\max_{i \notin \mathcal{I}} \theta_i^*$ with certainty, and $\alpha_{n,i} \to 0$ for every $i \notin \mathcal{I}$ except for I. This means there is a time $\tau' > \tau$ such that for $n \geq \tau'$

$$\alpha_{n,i} > \alpha' \text{ if } i \in \mathcal{I}$$

$$\alpha_{n,i} \leq \alpha' \text{ if } i \notin \mathcal{I} \text{ and } i \neq I.$$

When this occurs at least one of the two designs with highest probability $\alpha_{n,i}$ of being optimal must be in the set $\mathcal{I}$, which implies designs in $\mathcal{I}$ receive infinite measurement effort, yielding a contradiction.

Since $\sum_1^\infty \psi_{n,i} = \infty$ for all i, Lemma EC.12 implies $\alpha_{n,I^*} \to 1$. Therefore, there is a finite time τ such that $\hat{I}_n \triangleq \arg\max_i \alpha_{n,i} = I^*$ for all $n \geq \tau$. By the definition of the algorithm $\psi_{n,\hat{I}_n} = \beta$, and so $\psi_{n,I^*} = \beta$ for all $n \geq \tau$. We conclude that $\overline{\psi}_{n,I^*} \to \beta$.

Step 2: Show (EC.20) *holds.* As argued above, for each sample path there is a finite time $\tau < \infty$ such that for all $n \geq \tau$, $\hat{I}_n = I^*$ and therefore $\hat{J}_n = \arg\max_{i \neq I^*} \alpha_{n,i}$. By Lemma EC.11, one can choose $\tau' \geq \tau$ such that for all $n \geq \tau'$,

$$\overline{\psi}_{n,j} \geq \psi_j^\beta + \delta \implies \alpha_{n,j} < \max_{i \neq I^*} \alpha_{n,i}$$

and therefore by definition $\hat{J}_n \neq j$. This concludes the proof, as it shows that for each sample path there is a finite time τ' after which TTPS never allocates any measurement effort to design j when $\overline{\psi}_{n,j} \geq \psi_j^\beta + \delta$. $\square$

EC.8.2.2. Top-Two Value Sampling Recall that top-two value sampling sets $\psi_{n,\hat{I}_n} = \beta$ and $\psi_{n,\hat{J}_n} = 1 - \beta$ where $\hat{I}_n = \arg\max_i V_{n,i}$ and $\hat{J}_n = \arg\max_{j \neq \hat{I}_n} V_{n,i}$ are the two designs with the highest posterior value.

Proof for TTVS. *Step 1: Show* $\overline{\psi}_{n,I^*} \to \beta$. The proof is essentially identical to that for TTPS.To begin, we show $\sum_{n\in\mathbb{N}} \psi_{n,i} = \infty$ for each design i. Suppose otherwise. Let $\mathcal{I} = \{i \in \{1,..,k\} : \sum_1^\infty \psi_{n,i} < \infty\}$ be the set of designs to which finite measurement effort is allocated. Proceeding by contradiction, suppose $\mathcal{I}$ is nonempty. By Lemma EC.12, there is a time τ and some $v > 0$ such that $V_{n,i} > v$ for all $n \geq \tau$ and $i \in \mathcal{I}$. However, because of the assumption that $\theta_i^* \neq \theta_j^*$, for $i \neq j$, $I = \arg\max_{i \notin \mathcal{I}} \theta_i^*$ is unique6. . By (EC.21), the algorithm identifies $\arg\max_{i\notin\mathcal{I}} \theta_i^*$ with certainty, and $V_{n,i} \to 0$ for every $i \notin \mathcal{I}$ except for I. Then there is a time $\tau' > \tau$ such that for $n \geq \tau'$

$$V_{n,i} > v \text{ if } i \in \mathcal{I}$$

$$V_{n,i} \leq v \text{ if } i \notin \mathcal{I} \text{ and } i \neq I^*.$$

When this occurs at least one of the two designs with highest value $V_{n,i}$ must be in the set $\mathcal{I}$, which implies designs in $\mathcal{I}$ receive infinite measurement effort, yielding a contradiction.

Since $\sum_1^\infty \psi_{n,i} = \infty$ for all i, Lemma EC.12 implies $V_{n,I^*} \to v_{I^*}(\boldsymbol{\theta}^*) > 0$ and $V_{n,i} \to 0$ for all $i \neq I^*$. Therefore, there is a finite time τ such that $\arg\max_i V_{n,i} = I^*$ for all $\tau \geq n$. By the definition of the algorithm $\arg\max_i V_{n,i}$ is sampled with probability β, and so $\psi_{n,I^*} = \beta$ for all $n \geq \tau$. We conclude that $\overline{\psi}_{n,I^*} \to \beta$.

Step 2: Show (EC.20) *holds.* Again, the proof is essentially identical to that for TTPS. As argued above, for each sample path there is a finite time $\tau < \infty$ such that for all $n \geq \tau$, $\hat{I}_n = I^*$ and therefore $\hat{J}_n = \arg\max_{i\neq I^*} V_{n,i}$. By Lemma 4, $V_{n,i} \doteq \alpha_{n,i}$. Combining this with Lemma EC.11 shows one can choose $\tau' \geq \tau$ such that for all $n \geq \tau'$,

$$\overline{\psi}_{n,j} \geq \psi_j^\beta + \delta \implies V_{n,j} < \max_{i \neq I^*} V_{n,i}$$

and therefore by definition $\hat{J}_n \neq j$. This concludes the proof, as it shows that for each sample path there is a finite time τ' after which TTVS never allocates any measurement effort to design $j \neq I^*$ when $\overline{\psi}_{n,j} \geq \psi_j^\beta + \delta$. $\square$

EC.9. Results on Adaptive Tuning

PROPOSITION 3. *Suppose TTTS, TTVS, TTPS are applied with an adaptive sequence of tuning parameters $(\beta_n : n \in \mathbb{N})$ where for each n, β_n is $\mathcal{F}_{n-1}$ measurable. Then, with probability 1, on any sample path on which $\beta_n \to \beta^*$,*

$$\Pi_n(\Theta^c_{I^*}) \doteq e^{-n\Gamma^*}.$$

Proof. *Step 1: Show $\sum_{n\in\mathbb{N}} \psi_{n,i} = \infty$ for each design i.* The proof follows identically to the case of fixed β. For example, consider the case of TTTS and, proceeding by contradiction, suppose $\sum_{n\in\mathbb{N}} \psi_{n,i} < \infty$ on some sample path. Under TTTS, $\psi_{n,i} \geq \beta_n \alpha_{n,i}$, where $\beta_n \to \beta^* > 0$. By Lemma EC.12, if $i \in \mathcal{I}$ then $\liminf_{n\to\infty} \alpha_{n,i} > 0$. Together this shows $\liminf_{n\in\mathbb{N}} \psi_{n,i} > 0$. This implies $\sum_{n\in\mathbb{N}} \psi_{n,i} = \infty$, a contradiction. Proofs for TTPS and TTVS also follows as before, and are omitted

Step 2: Show that $\overline{\psi}_{n,I^*} \to \beta^*$

It is sufficient to show that $\psi_{n,I^*} - \beta_n \to 0$. This would imply $\frac{1}{n}\sum_{\ell=1}^{n} (\psi_{\ell,I^*} - \beta_\ell) \to 0$, which since $\beta_n \to \beta^*$, implies $\frac{1}{n}\sum_{\ell=1}^{n} \psi_{\ell,I^*} \to \beta^*$ as desired.

Now, since $\sum_{n\in\mathbb{N}} \psi_{n,i} = \infty$ for all arms i, Lemma EC.12 implies

$$V_{n,i} \to \begin{cases} 0 & \text{if } i \neq I^* \\ v_{I^*}(\boldsymbol{\theta}^*) > 0 & \text{if } i = I^* \end{cases} \qquad \text{and} \qquad \alpha_{n,i} \to \begin{cases} 0 & \text{if } i \neq I^* \\ 1 & \text{if } i = I^*. \end{cases}$$

For top-two probability sampling, this implies there exists a time after which $\arg\max_i \alpha_{n,i} = I^*$ and hence $\psi_{n,I^*} = \beta_n$ for all n sufficiently large. For top-two value sampling, the same result applies, since there exists a time after which $\arg\max_i V_{n,i} = I^*$. For top-two Thompson sampling (See Subsection 3.4),

$$\psi_{n,i} = \alpha_{n,i}\left(\beta_n + (1-\beta_n)\sum_{j\neq i} \frac{\alpha_{n,j}}{1-\alpha_{n,j}}\right).$$

from which we conclude $\psi_{n,I^*} - \beta_n \to 0$ as $\alpha_{n,I^*} \to 1$.

Step 3: Show sufficient condition for optimality in (EC.20). By Lemma EC.10, it is enough to show

$$\sum_{n\in\mathbb{N}} \psi_{n,j}\mathbf{1}(\overline{\psi}_{n,j} \geq \psi_j^{\beta^*} + \delta) < \infty \qquad \forall\, j \neq I^*,\, \delta > 0,$$

For each proposed algorithm, a proof of the corresponding result was given in Step 2 of Subsection EC.8.2, but for the case of arbitrary $\beta \in (0,1)$. Since $\overline{\psi}_{n,I^*} \to \beta^*$, for each of proposed algorithm the proof of this follows line by line as before, but replacing β with β^* everywhere it occurs. □

LEMMA 1. *Under TTTS, TTPS, or TTVS with an adaptive sequence of tuning parameters $(\beta_n : n \in \mathbb{N})$ adjusted according to Algorithm 3, $\beta_n \to \beta^*$ almost surely. Therefore $\Pi_n(\Theta_{I^*}^c) \doteq e^{-n\Gamma^*}$.*

Proof. First, let us define some notation. Let $\hat{\boldsymbol{\theta}}_n = \int_{\boldsymbol{\theta}\in\Theta} \boldsymbol{\theta}\pi_n(\boldsymbol{\theta})d\boldsymbol{\theta}$ denote the posterior mean at time n. Recall that β_n denotes the tuning parameter used by the top-two algorithm at time n and this is updated only at certain time periods. Define

$$\ell_n = \max\{\ell \in \mathbb{N} : \min_{i\in\{1,\ldots k\}} S_{n,i} \geq \kappa^\ell\} = \left\lfloor \log_\kappa \left(\min_i S_{n,i} \right) \right\rfloor$$

to be the number of time periods in which an update to β_n has been attempted. Now, the proof proceeds in two steps.

Step 1: Show $\sum_{n\in\mathbb{N}} \psi_{n,i} = \infty$ for all $i \in \{1,\ldots,k\}$ almost surely. We give a short sketch of a proof by contradiction. Suppose otherwise. Then by Corollary EC.1, we know that there is some arm $i \in \{1,\ldots k\}$ with $\lim_{n\to\infty} S_{n,i} < \infty$. This in turn implies $\sup_{n\in\mathbb{N}} \ell_n < \infty$, so on this sample path there exists a time N with $\beta_n = \beta_N$ for all $n \geq N$. Let us consider the sample path from time N onwards. We have concluded that there is an infinite period of times $n \in \{N_1, N_2, \ldots\}$ over which (1) top-two sampling is applied with a constant parameter $\beta_n = \beta_N$ with initial beliefs π_N over the hyper-rectangle $\Theta = (\underline{\theta}, \overline{\theta})^k$ and (2) there exists an arm i with $\sum_{n=N}^{\infty} \psi_{n,i} < \infty$. We know by Proposition 8 and its proof that the set of such sample paths has measure zero.

Step 2: Show that therefore $\beta_n \to \beta^*$. We first show the consistency of the posterior mean

$\hat{\boldsymbol{\theta}}_n = \int_{\boldsymbol{\theta}\in\Theta} d\pi_n(\boldsymbol{\theta})$. Since $\sum_{n\in\mathbb{N}} \psi_{n,i} = \infty$ for all i, Proposition 4 implies that Π_n converges to a point mass at $\boldsymbol{\theta}^*$, in the sense that for any open set $\tilde{\Theta}$ containing $\boldsymbol{\theta}^*$, $\Pi_n\left(\tilde{\Theta}\right) \to 1$ almost surely. Since Θ is compact, this implies $\int_{\boldsymbol{\theta}\in\Theta} d\pi_n(\boldsymbol{\theta}) \to \boldsymbol{\theta}^*$. (Formally, one can justify the exchange of limit an integral by applying the bounded convergence theorem to each element $\hat{\theta}_{n,i}$.) Now, because the function $f(\boldsymbol{\psi};\boldsymbol{\theta}) := \min_{\boldsymbol{\theta}'\in\Theta^c_{\hat{I}}} D_{\psi}(\boldsymbol{\theta}||\boldsymbol{\theta}')$ is continuous in both arguments, the correspondence $\boldsymbol{\theta} \mapsto \arg\max_{\psi} f(\boldsymbol{\psi},\boldsymbol{\theta})$ is upper hemi-continuous at any $\boldsymbol{\theta}$. Since $\psi^*(\boldsymbol{\theta}^*) = \arg\max_{\boldsymbol{\psi}} f(\boldsymbol{\psi},\boldsymbol{\theta}^*)$ is unique, we know $\psi^*(\boldsymbol{\theta})$ is continuous in a neighborhood of $\boldsymbol{\theta}^*$. This implies $\psi^*(\hat{\boldsymbol{\theta}}_n) \to \psi^*(\boldsymbol{\theta}^*)$. Let $\hat{I}_n = \arg\max_i \hat{\theta}_{n,i}$ denote the estimate of the best arm, and note that $\hat{I}_n \to I^*$ almost surely. This gives $\boldsymbol{\psi}^*(\hat{\boldsymbol{\theta}}_n)_{\hat{I}_n} \to \boldsymbol{\psi}^*(\boldsymbol{\theta}^*)_{I^*} = \beta^*$. Since the parameter β_n is updated an infinite number of times, this implies $\beta_n \to \beta^*$. $\square$