Electronic Companion to Top-Two Thompson Sampling for Contextual Selection Problems

Appendix A: Posterior large deviations principle. In this section, we prove Theorem 1 which extends the posterior large deviations principle for general identifiable distributions. Since Assumption 2 and Assumption 2′ only differ from the identifiable assumption, which is not used in the proof of Theorem 1, we suppose Assumption 2 is valid without loss of generality.

Proof of Theorem 1. We first establish a uniform approximation to $W_T(\cdot)$ by the overall KL divergence, namely,

$$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \left| W_T(\boldsymbol{\theta}) - D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta}) \right| \to 0, \quad a.s. \tag{EC.1}$$

Recall that $W_T(\boldsymbol{\theta}) = \sum_{d\in\mathcal{D}} \frac{1}{T}\sum_{t=1}^{T} \xi_t(d)\lambda(\theta_d^*, \theta_d; Y_d^{(t)})$ is the weighted log-likelihood ratio and $D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta}) = \sum_{d\in\mathcal{D}} \bar{\psi}_T(d) D(\theta_d^*\|\theta_d)$ is the weighted KL divergence. Notice that (EC.1) does not hold directly from the GC class assumption since $Y_d^{(t)}$ is not independent of the weights $\xi_t(d)$.

Fix a design $d\in\mathcal{D}$, we define a sequence of strictly increasing stopping times by recursion

$$\tau_1 = \min\{t\in\mathbb{N} : \xi_t(d) = 1\}, \quad \text{and} \quad \tau_{k+1} = \min\{\tau_k < t\in\mathbb{N} : \xi_t(d) = 1\}, \quad k = 1, 2, \dots, \tag{EC.2}$$

with the convention $\min\emptyset = \infty$ for empty set $\emptyset$. We note that these stopping times are well defined on the extended filtration $\{\mathcal{E}_t' := \sigma(\mathcal{E}_{t-1} \bigcup \sigma(\xi_1(d), \dots, \xi_t(d)))\}$ rather than $\{\mathcal{E}_{t-1}\}$. Then,

$$\frac{1}{T}\sum_{t=1}^{T} \xi_t(d)\lambda\left(\theta_d^*, \theta_d; Y_d^{(t)}\right) = \frac{1}{T}\sum_{k=1}^{T} \mathbf{1}\{\tau_k \le T\}\lambda\left(\theta_d^*, \theta_d; Y_d^{(\tau_k)}\right). \tag{EC.3}$$

Let $Z_d^{(t)}$ and ${Z'}_d^{(t)}$ be two independent copies of Y_d that are also independent of anything else. We investigate the convergence of the right-hand side of Equation (EC.3) by its modification with $Y_d^{(\tau_k)}$ replaced by $Z_d^{(k)}$. Denote $a_k = \mathbf{1}\{\tau_k < \infty\}\lambda\left(\theta_d^*, \theta_d; Y_d^{(\tau_k)}\right) + \mathbf{1}\{\tau_k = \infty\}\lambda\left(\theta_d^*, \theta_d; {Z'}_d^{(k)}\right)$ and $b_k = \lambda\left(\theta_d^*, \theta_d; Z_d^{(k)}\right)$. Note that a_k is well-defined even when $\tau_k = \infty$. Then we will show that $\{a_k\}$ is a process identically distributed as $\{b_k\}$. For any Borel measurable sets $A_1, A_2, \dots$, the following equality holds for all k:

$$\begin{aligned}
\mathbb{E}\left[\mathbf{1}\{a_k \in A_k\} \middle| \mathcal{E}'_{\tau_k}\right] =& P\left(\lambda\left(\theta_d^*, \theta_d; Y_d^{(\tau_k)}\right) \in A_k \middle| \tau_k < \infty, \mathcal{E}'_{\tau_{k-1}}\right) P(\tau_k < \infty | \mathcal{E}'_{\tau_{k-1}}) \\
&+ P\left(\lambda\left(\theta_d^*, \theta_d; {Z'}_d^{(k)}\right) \in A_k \middle| \tau_k < \infty, \mathcal{E}'_{\tau_{k-1}}\right) P(\tau_k = \infty | \mathcal{E}'_{\tau_{k-1}}) \\
=& P\left(\lambda\left(\theta_d^*, \theta_d; Y_d\right) \in A_k\right) \left(P(\tau_k < \infty | \mathcal{E}'_{\tau_{k-1}}) + P(\tau_k = \infty | \mathcal{E}'_{\tau_{k-1}})\right) \\
=& P\left(\lambda\left(\theta_d^*, \theta_d; Y_d\right) \in A_k\right),
\end{aligned} \tag{EC.4}$$

where $\{\mathcal{E}'_{\tau_k} : k \in \mathbb{N}\}$ is a sequence of stopping time σ-algebra defined as $\mathcal{E}'_{\tau_k} := \{A : A \bigcap \{\tau_k \leq n\} \in \mathcal{E}'_n,\ \forall n \geq 1\}$, $\forall k \in \mathbb{N}$. Therein, to see the validity of the second equality, note that it follows from the definition of ${Z'_d}^{(k)}$ and the following equality:

$$\begin{aligned}
&P\left(\lambda\left(\theta_d^*, \theta_d; Y_d^{(\tau_k)}\right) \in A_k \middle| \tau_k < \infty, \mathcal{E}'_{\tau_{k-1}}\right) \\
=&\sum_{m=1}^{\infty} P\left(\lambda\left(\theta_d^*, \theta_d; Y_d^{(\tau_k)}\right) \in A_k, \tau_{k-1} = m \middle| \tau_k < \infty, \mathcal{E}'_m\right) \\
=&\sum_{m=1}^{\infty} \sum_{n=m+1}^{\infty} P\left(\tau_k = n | \tau_k < \infty, \mathcal{E}'_m\right) P\left(\lambda\left(\theta_d^*, \theta_d; Y_d^{(n)}\right) \in A_k, \tau_{k-1} = m \middle| \tau_k = n, \mathcal{E}'_m\right) \\
\overset{(i)}{=}&\sum_{m=1}^{\infty} \sum_{n=m+1}^{\infty} P\left(\tau_k = n | \tau_k < \infty, \mathcal{E}'_m\right) P\left(\lambda\left(\theta_d^*, \theta_d; Y_d^{(n)}\right) \in A_k \middle| \tau_k = n, \mathcal{E}'_m\right) \mathbf{1}\{\tau_{k-1} = m\} \\
\overset{(ii)}{=}&\sum_{m=1}^{\infty} \sum_{n=m+1}^{\infty} P\left(\tau_k = n, \tau_{k-1} = m | \tau_k < \infty, \mathcal{E}'_m\right) \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}\{\lambda(\theta_d^*, \theta_d; Y_d^{(n)}) \in A_k\} \middle| \mathcal{E}'_n\right] \middle| \tau_k = n, \mathcal{E}'_m\right] \\
=&\sum_{m=1}^{\infty} \sum_{n=m+1}^{\infty} P\left(\tau_k = n, \tau_{k-1} = m | \tau_k < \infty, \mathcal{E}'_m\right) \mathbb{E}\left[P\left(\lambda(\theta_d^*, \theta_d; Y_d) \in A_k\right) | \tau_k = n, \mathcal{E}'_m\right] \\
=&\sum_{m=1}^{\infty} \sum_{n=m+1}^{\infty} P\left(\tau_k = n, \tau_{k-1} = m | \tau_k < \infty, \mathcal{E}'_m\right) P\left(\lambda(\theta_d^*, \theta_d; Y_d) \in A_k\right) \\
=&P\left(\lambda(\theta_d^*, \theta_d; Y_d) \in A_k\right),
\end{aligned}$$

where the first equality follows from the definition of $\mathcal{E}'_{\tau_{k-1}}$. Equality (i) is due to the fact $\{\tau_{k-1} = m\} \in \mathcal{E}'_m$. Equality (ii) follows from the law of iterated expectation and the fact that $P\left(\tau_k = n | \tau_k < \infty, \mathcal{E}'_m\right) \mathbf{1}\{\tau_{k-1} = m\} = P\left(\tau_k = n, \tau_{k-1} = m | \tau_k < \infty, \mathcal{E}'_m\right)$, which is implied by $\{\tau_k = n\} \in \mathcal{E}'_n$.

Therefore, we see that

$$\begin{aligned}
&\mathbb{P}((a_1, a_2, \ldots, a_T) \in A_1 \times A_2 \times \cdots \times A_T) \\
=\ &\mathbb{E}\left[\prod_{k=1}^{T} \mathbf{1}\{a_k \in A_k\}\right] \\
\overset{(i)}{=}\ &\mathbb{E}\left[\mathbb{E}\left[\prod_{k=1}^{T} \mathbf{1}\{a_k \in A_k\} \middle| \mathcal{E}'_{\tau_T}, {Z'_d}^{(1)}, \ldots, {Z'_d}^{(T-1)}\right]\right] \\
\overset{(ii)}{=}\ &\mathbb{E}\left[\prod_{k=1}^{T-1} \mathbf{1}\{a_k \in A_k\} \cdot \mathbb{E}\left[\mathbf{1}\{a_T \in A_T\} \middle| \mathcal{E}'_{\tau_T}, {Z'_d}^{(1)}, \ldots, {Z'_d}^{(T-1)}\right]\right] \\
\overset{(iii)}{=}\ &\mathbb{E}\left[\prod_{k=1}^{T-1} \mathbf{1}\{a_k \in A_k\}\right] \cdot \mathbb{P}\left(\lambda\left(\theta_d^*, \theta_d; Y_d\right) \in A_T\right) \\
\overset{(iv)}{=}\ &\prod_{k=1}^{T} \mathbb{P}\left(\lambda\left(\theta_d^*, \theta_d; Y_d\right) \in A_k\right) \\
\overset{(v)}{=}\ &\mathbb{P}((b_1, b_2, \ldots, b_T) \in A_1 \times A_2 \times \cdots \times A_T).
\end{aligned}$$

Therein, equality (i) follows from the law of iterated expectation. Equality (ii) holds because a_k is measurable in the sigma field generated by $\mathcal{E}'_{\tau_T}$ and ${Z'_d}^{(1)}, \ldots, {Z'_d}^{(T-1)}$. Recall that a_T depends on ${Z'_d}^{(T)}$ but is independent of ${Z'_d}^{(1)}, \ldots, {Z'_d}^{(T-1)}$. Therefore, we have $\mathbb{E}\left[\mathbf{1}\{a_T \in A_T\} \middle| \mathcal{E}'_{\tau_T}, {Z'_d}^{(1)}, \ldots, {Z'_d}^{(T-1)}\right] = \mathbb{E}\left[\mathbf{1}\{a_T \in A_T\} \middle| \mathcal{E}'_{\tau_T}\right]$, and then equality (iii) follows from Equation (EC.4). Finally, equality (iv) follows by induction and equality (v) follows by the definition of $b_k = \lambda\left(\theta_d^*, \theta_d; Z_d^{(k)}\right)$, where ${Z'_d}^{(k)}$ are independent of each other.

Thus $(a_1,a_2,\dots,a_T)$ and $(b_1,b_2,\dots,b_T)$ have the same probability on any rectangular set in $\mathbb{R}^T$ for any T. Therefore, the processes $\{a_k\}$ and $\{b_k\}$ are identically distributed by Kolmogorov's extension theorem. Moreover, let B be the event where the design d is sampled infinitely often, that is, $B:=\{\tau_k<\infty,\ \forall k\in\mathbb{N}\}$. Then we see that

$$\begin{aligned}
&\mathbb{P}\left(\left\{\limsup_{T\to\infty}\sup_{\theta_d\in\Theta}\left|\frac{1}{T}\sum_{k=1}^{T}\mathbf{1}\{\tau_k\le T\}\left(\lambda\left(\theta_d^*,\theta_d;Y_d^{(\tau_k)}\right)-D(\theta_d^*\|\theta_d)\right)\right|>0\right\}\bigcap B\right)\\
\le\ &\mathbb{P}\left(\left\{\limsup_{T\to\infty}\sup_{\theta_d\in\Theta}\left|\sum_{k=1}^{T}\frac{\mathbf{1}\{\tau_k\le T\}}{\sum_{l=1}^{T}\mathbf{1}\{\tau_l\le T\}}\left(\lambda\left(\theta_d^*,\theta_d;Y_d^{(\tau_k)}\right)-D(\theta_d^*\|\theta_d)\right)\right|>0\right\}\bigcap B\right)\\
\overset{(i)}{=}\ &\mathbb{P}\left(\left\{\limsup_{T\to\infty}\sup_{\theta_d\in\Theta}\left|\frac{1}{\max\{k\ge1:\tau_k\le T\}}\sum_{k=1}^{\max\{k\ge1:\tau_k\le T\}}(a_k-D(\theta_d^*\|\theta_d))\right|>0\right\}\bigcap B\right)\\
\overset{(ii)}{\le}\ &\mathbb{P}\left(\left\{\limsup_{T'\to\infty}\sup_{\theta_d\in\Theta}\left|\frac{1}{T'}\sum_{k=1}^{T'}(a_k-D(\theta_d^*\|\theta_d))\right|>0\right\}\right)\\
\overset{(iii)}{=}\ &\mathbb{P}\left(\left\{\limsup_{T'\to\infty}\sup_{\theta_d\in\Theta}\left|\frac{1}{T'}\sum_{k=1}^{T'}(b_k-D(\theta_d^*\|\theta_d))\right|>0\right\}\right)=\ 0,
\end{aligned}\tag{EC.5}$$

where equality (i) follows from the fact that τ_k is strictly increasing, inequality (ii) follows from $\lim_{T\to\infty}\max\{k\ge1:\tau_k\le T\}=\infty$ in B, equality (iii) follows from the fact that $\{a_k\}$ and $\{b_k\}$ are identically distributed, and the last equality follows from the GC class assumption.

Moreover, note that for any $T\ge1$, there are only a finite number of stopping times among τ_k's that are finite for each sample path in B^C. Therefore, we have

$$\begin{aligned}
&\mathbb{P}\left(\left\{\limsup_{T\to\infty}\sup_{\theta_d\in\Theta}\left|\frac{1}{T}\sum_{k=1}^{\infty}\mathbf{1}\{\tau_k\le T\}\left(\lambda\left(\theta_d^*,\theta_d;Z_d^{(k)}\right)-D(\theta_d^*\|\theta_d)\right)\right|>0\right\}\bigcap B^C\right)\\
\le&\mathbb{P}\left(\left\{\limsup_{T\to\infty}\frac{1}{T}\sum_{k=1}^{\infty}\mathbf{1}\{\tau_k\le\infty\}\sup_{\theta_d\in\Theta}\left|\lambda\left(\theta_d^*,\theta_d;Z_d^{(k)}\right)-D(\theta_d^*\|\theta_d)\right|>0\right\}\bigcap B^C\right)=0,
\end{aligned}\tag{EC.6}$$

since for all sample path in B^C, $\sup_{\theta_d\in\Theta}|\lambda(\theta_d^*,\theta_d;Z_d^{(k)})-D(\theta_d^*\|\theta_d)|$ is bounded due to the compactness assumption. Combining Equations (EC.3)-(EC.6),

$$\sup_{\theta_d\in\Theta}\left|\frac{1}{T}\sum_{t=1}^{T}\xi_t(d)\lambda\left(\theta_d^*,\theta_d;Y_d^{(t)}\right)-\frac{1}{T}\sum_{t=1}^{T}\xi_t(d)D(\theta_d^*\|\theta_d)\right|\to0,\quad a.s.\tag{EC.7}$$

Note that $\sum_{t=1}^{T}(\xi_t(d)-\psi_t(d))$ is a bounded martingale difference sequence with respect to the filtration $\mathcal{E}_t'$, and the KL divergence $D(\theta_d^*\|\theta_d)$ is universally bounded. Therefore, the law of large numbers of martingale difference sequences [12] implies

$$\sup_{\theta_d\in\Theta}\left|\frac{1}{T}\sum_{t=1}^{T}\xi_t(d)D(\theta_d^*\|\theta_d)-\frac{1}{T}\sum_{t=1}^{T}\psi_t(d)D(\theta_d^*\|\theta_d)\right|\to0,\quad a.s.\tag{EC.8}$$

To partially conclude, Equation (EC.1) follows from (EC.7) and (EC.8), and the triangle inequality. An immediate consequence of (EC.1) is

$$\lim_{T\to\infty}-\frac{1}{T}\ln\int_{\tilde{\boldsymbol{\Theta}}}\exp\{-TW_T(\boldsymbol{\theta})\}d\boldsymbol{\theta}+\frac{1}{T}\ln\int_{\tilde{\boldsymbol{\Theta}}}\exp\{-TD_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})\}d\boldsymbol{\theta}=0,\quad a.s.\tag{EC.9}$$

To see this, note that

$$\begin{aligned}
&-\frac{1}{T}\ln\int_{\tilde{\boldsymbol{\Theta}}}\exp\{-TW_T(\boldsymbol{\theta})\}d\boldsymbol{\theta}\\
=&-\frac{1}{T}\ln\int_{\tilde{\boldsymbol{\Theta}}}\exp\{-T(D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})+W_T(\boldsymbol{\theta})-D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta}))\}d\boldsymbol{\theta}\\
\le&-\frac{1}{T}\ln\int_{\tilde{\boldsymbol{\Theta}}}\exp\{-TD_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})\}d\boldsymbol{\theta}+\sup_{\boldsymbol{\theta}\in\tilde{\boldsymbol{\Theta}}}(W_T(\boldsymbol{\theta})-D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})),
\end{aligned}$$

and similarly,

$$\begin{aligned}
&-\frac{1}{T}\ln\int_{\tilde{\boldsymbol{\Theta}}}\exp\{-TW_T(\boldsymbol{\theta})\}d\boldsymbol{\theta}\\
\ge&-\frac{1}{T}\ln\int_{\tilde{\boldsymbol{\Theta}}}\exp\{-TD_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})\}d\boldsymbol{\theta}-\inf_{\boldsymbol{\theta}\in\tilde{\boldsymbol{\Theta}}}(W_T(\boldsymbol{\theta})-D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})).
\end{aligned}$$

Secondly, using Laplace approximations, we show that for any open set $\tilde{\boldsymbol{\Theta}}\subseteq\boldsymbol{\Theta}$,

$$\lim_{T\to\infty}-\frac{1}{T}\ln\int_{\tilde{\boldsymbol{\Theta}}}\exp\{-TD_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})\}d\boldsymbol{\theta}-\inf_{\boldsymbol{\theta}\in\tilde{\boldsymbol{\Theta}}}D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})=0. \tag{EC.10}$$

The argument is similar to that used in [45]. The left-hand side is apparently non-negative. For the other direction, let $r:=\sup_{\theta\in\Theta}\|\partial D(\theta^*\|\theta)/\partial\theta\|$, which is finite since Θ is compact, and $D(\theta^*\|\theta)$ is continuously differentiable. For any $r\varepsilon>0$, one can choose a finite ε-cover O of $\tilde{\boldsymbol{\Theta}}$ such that each $o\in O$ intersects $\tilde{\boldsymbol{\Theta}}$. Let $\boldsymbol{\theta}_T\in\boldsymbol{\Theta}$ that achieves the infimum $\inf_{\boldsymbol{\theta}\in\tilde{\boldsymbol{\Theta}}}D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})$, which possibly does not lie in $\tilde{\theta}$, and $o_T\in O$ be any set whose closure contains $\boldsymbol{\theta}_T$. Then,

$$\begin{aligned}
-\frac{1}{T}\ln\int_{\tilde{\boldsymbol{\Theta}}}\exp\{-TD_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})\}d\boldsymbol{\theta}&\le-\frac{1}{T}\ln\int_{\tilde{\boldsymbol{\Theta}}\bigcap o_T}\exp\{-TD_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})\}d\boldsymbol{\theta}\\
&\le-\frac{1}{T}\ln\int_{\tilde{\boldsymbol{\Theta}}\bigcap o_T}\exp\{-T(D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta}_T)+r\varepsilon)\}d\boldsymbol{\theta}\\
&=-\frac{1}{T}\ln\left(\int_{\tilde{\boldsymbol{\Theta}}\bigcap o_T}d\boldsymbol{\theta}\cdot\exp\{-T\cdot\inf_{\boldsymbol{\theta}\in\tilde{\boldsymbol{\Theta}}}D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})\}\cdot\exp\{-T\cdot r\varepsilon\}\right)\\
&=-\frac{1}{T}\ln\int_{\tilde{\boldsymbol{\Theta}}\bigcap o_T}d\boldsymbol{\theta}+\inf_{\boldsymbol{\theta}\in\tilde{\boldsymbol{\Theta}}}D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})+r\varepsilon,
\end{aligned}$$

where the second inequality follows from $\left|D(\boldsymbol{\theta}^*\|\boldsymbol{\theta})-D(\boldsymbol{\theta}^*\|\boldsymbol{\theta}_T)\right|\le r\|\boldsymbol{\theta}-\boldsymbol{\theta}_T\|\le r\varepsilon$. The third equality follows from the fact that $\exp\{-T(D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta}_T)+r\varepsilon)\}=\exp\{-T\cdot D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta}_T)\}\cdot\exp\{-T\cdot r\varepsilon\}$ and that $D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta}_T)=\inf_{\boldsymbol{\theta}\in\tilde{\boldsymbol{\Theta}}}D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})$ by the definition of $\boldsymbol{\theta}_T$. Sending T to infinity and choosing ε to be arbitrarily small yields Equation (EC.10).

Equations (EC.9) and (EC.10) combine to yield

$$\lim_{T\to\infty}-\frac{1}{T}\ln\int_{\tilde{\boldsymbol{\Theta}}}\exp\{-TW_T(\boldsymbol{\theta})\}d\boldsymbol{\theta}-\inf_{\boldsymbol{\theta}\in\tilde{\boldsymbol{\Theta}}}D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})=0,\quad a.s. \tag{EC.11}$$

Therefore, it follows from (EC.11) by choosing $\tilde{\boldsymbol{\Theta}}=\mathrm{int}(\boldsymbol{\Theta})$ that

$$\lim_{T\to\infty}-\frac{1}{T}\ln\int_{\mathrm{int}(\boldsymbol{\Theta})}\exp\{-TW_T(\boldsymbol{\theta})\}d\boldsymbol{\theta}-\inf_{\boldsymbol{\theta}\in\mathrm{int}(\boldsymbol{\Theta})}D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})=0,\quad a.s. \tag{EC.12}$$

Note that $\ln\Pi_T(\tilde{\boldsymbol{\Theta}})=\ln\int_{\tilde{\boldsymbol{\Theta}}}\exp\{-TW_T(\boldsymbol{\theta})\}d\boldsymbol{\theta}-\ln\int_{\mathrm{int}(\boldsymbol{\Theta})}\exp\{-TW_T(\boldsymbol{\theta})\}d\boldsymbol{\theta}$, where $\mathrm{int}(\cdot)$ indicates the interior part of a set. Then the first statement in Theorem 1 follows from (EC.11) and (EC.12) and $\inf_{\boldsymbol{\theta}\in\mathrm{int}(\boldsymbol{\Theta})}D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})=\inf_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})=D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta}^*)=0$.

Finally, we prove the second statement. Fix a context c and a design $d \in \mathcal{D}_c$. By checking the proof of (EC.7), we see that if $\sum_{t=1}^{T}\sum_{d'\in\mathcal{D}_c}\xi_t(d') \to \infty$, then the following stronger version of (EC.7) holds:

$$\sup_{\theta_d\in\Theta}\left|\frac{\sum_{t=1}^{T}\xi_t(d)\lambda\left(\theta_d^*,\theta_d;Y_d^{(t)}\right)}{\sum_{t=1}^{T}\sum_{d'\in\mathcal{D}_c}\xi_t(d')} - \frac{\sum_{t=1}^{T}\xi_t(d)D(\theta_d^*\|\theta_d)}{\sum_{t=1}^{T}\sum_{d'\in\mathcal{D}_c}\xi_t(d')}\right| \to 0. \tag{EC.13}$$

Actually, the left-hand side of (EC.13) coincides with that of (EC.7) with T substituted by $\sum_{t=1}^{T}\sum_{d'}\xi_t(d')$, i.e., the number of samples allocated to context c. The second probability of (EC.5) is indeed equal to

$$\mathbb{P}\left(\left\{\limsup_{T\to\infty}\sup_{\theta_d\in\Theta}\left|\frac{\sum_{t=1}^{T}\xi_t(d)\lambda\left(\theta_d^*,\theta_d;Y_d^{(t)}\right)}{\sum_{t=1}^{T}\sum_{d'\in\mathcal{D}_c}\xi_t(d')} - \frac{\sum_{t=1}^{T}\xi_t(d)D(\theta_d^*\|\theta_d)}{\sum_{t=1}^{T}\sum_{d'\in\mathcal{D}_c}\xi_t(d')}\right| > 0\right\}\bigcap B\right) = 0,$$

and it follows from a similar argument to (EC.6) with T substituted by $\sum_{t=1}^{T}\sum_{d'}\xi_t(d')$ that

$$\mathbb{P}\left(\left\{\limsup_{T\to\infty}\sup_{\theta_d\in\Theta}\left|\frac{\sum_{t=1}^{T}\xi_t(d)\lambda\left(\theta_d^*,\theta_d;Y_d^{(t)}\right)}{\sum_{t=1}^{T}\sum_{d'\in\mathcal{D}_c}\xi_t(d')} - \frac{\sum_{t=1}^{T}\xi_t(d)D(\theta_d^*\|\theta_d)}{\sum_{t=1}^{T}\sum_{d'\in\mathcal{D}_c}\xi_t(d')}\right| > 0\right\}\bigcap B^C\right) = 0,$$

because $\sum_{t=1}^{T}\sum_{d'}\xi_t(d')$ tends to infinity according to the assumption. The two equations above imply (EC.13) holds.

By Levy's extension of the Borel-Cantelli Lemma [55], we see that with probability 1, if $T\bar{\alpha}_T(c) \to \infty$, then

$$\frac{\sum_{t=1}^{T}\sum_{d'\in\mathcal{D}_c}\xi_t(d')}{T\bar{\alpha}_T(c)} = \frac{\sum_{t=1}^{T}\sum_{d'\in\mathcal{D}_c}\xi_t(d')}{\sum_{t=1}^{T}\sum_{d'\in\mathcal{D}_c}\psi_t(d')} \to 1, \tag{EC.14}$$

and

$$\frac{\sum_{t=1}^{T}\xi_t(d)}{\sum_{t=1}^{T}\sum_{d'\in\mathcal{D}_c}\xi_t(d')} - \frac{\sum_{t=1}^{T}\psi_t(d)}{T\bar{\alpha}_T(c)} \to 0. \tag{EC.15}$$

In other words, the number of allocated simulation samples to d is in proportion to the sum of probabilities of sampling d in each step. Therefore,

$$\begin{aligned}
&\sup_{\theta_d\in\Theta}\left|\frac{\sum_{t=1}^{T}\xi_t(d)D(\theta_d^*\|\theta_d)}{\sum_{t=1}^{T}\sum_{d'\in\mathcal{D}_c}\xi_t(d')} - \frac{\sum_{t=1}^{T}\psi_t(d)D(\theta_d^*\|\theta_d)}{T\bar{\alpha}_T(c)}\right| \\
&= \sup_{\theta_d\in\Theta}|D(\theta_d^*\|\theta_d)|\cdot\left|\frac{\sum_{t=1}^{T}\xi_t(d)}{\sum_{t=1}^{T}\sum_{d'\in\mathcal{D}_c}\xi_t(d')} - \frac{\sum_{t=1}^{T}\psi_t(d)}{T\bar{\alpha}_T(c)}\right| \to 0.
\end{aligned} \tag{EC.16}$$

Combining (EC.13)-(EC.16) and the triangle inequality, we arrive at a counterpart of (EC.1)

$$\sup_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\left|\sum_{d\in\mathcal{D}_c}\frac{1}{T\bar{\alpha}_T(c)}\sum_{t=1}^{T}\xi_t(d)\lambda\left(\theta_d^*,\theta_d;Y_d^{(t)}\right) - D_{\bar{\beta}_T(c,\cdot)}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})\right| \to 0, \quad a.s.$$

Then the second statement follows by applying the Laplace approximation trick again. □

This proof necessitates Assumption 1 to control the speed at which the posterior density concentrates around the true parameter. To be specific, it allows us to control the integral in equation (EC.10) using a finite cover argument. It is possible to assume directly that for each $\varepsilon > 0$, there exists a finite cover of size $O(1)$ such that for any $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2$ in the same set belonging to this cover, $|D(\boldsymbol{\theta}^*\|\boldsymbol{\theta}_1) - D(\boldsymbol{\theta}^*\|\boldsymbol{\theta}_2)| \le \varepsilon$. This assumption is more fundamental but is harder to check than the second-order condition stated in Assumption 1. Moreover, the smoothness of KL divergence function is more of a consequence of the parameterization of the model. Reparameterization techniques may be used to change the smoothness of the KL divergence while keeping other aspects of the model unchanged. For example, suppose $f(x) = |x|$ is the KL divergence function for parameter x, which is not continuously differentiable. However, $g(y) = y^2$ is continuously differentiable with change of variable $y = \sqrt{|x|} \cdot (\mathbf{1}\{x > 0\} - \mathbf{1}\{x < 0\})$. Note that this assumption only serves for the purpose of the theoretical analysis and is not used in the algorithm. Therefore, another way to put Assumption 1 is to assume that there exists a bijection $T(\theta)$ such that $D(T(\theta^*)\|T(\theta))$ is continuously differentiable with respect to θ.

We briefly discuss the possibility to allow the parameter space Θ to expand as samples accumulate. Suppose that the decision maker considers an increasing set sequence $\{\Theta_T\}_{T\geq 1}$ where Θ_T is the parameter space for θ_d, $\forall d \in \mathcal{D}$, at time T. Correspondingly, the posterior probability measure $\Pi_T(\boldsymbol{\theta})$ will be modified by replacing the integral region $\boldsymbol{\Theta}$ with $\boldsymbol{\Theta}_T := \Theta_T^{|\mathcal{D}|}$. This is practically feasible if the focal Bayesian model assumes a conjugate prior distribution, because whatever the parameter space is, the posterior distribution is proportional to the likelihood function in the parameter space and can be easily calculated utilizing the conjugate property.

We assume alternative assumptions for the increasing parameter space in order for the posterior large deviation to hold.

Assumption EC.1. *Suppose $\Theta_1 \subseteq \Theta_2 \subseteq \cdots \subseteq \Theta_T \subseteq \cdots$. Assume that*

(1) The parameter space Θ_T, $\forall T \geq 1$ is a compact set in $\mathbb{R}^{k+1}$;
(2) The parametric family $\{p(\cdot|\theta) : \theta \in \bigcup_{T\geq 1} \Theta_T\}$ is pointwise continuous in θ;
(3) The parametric family $\{p(\cdot|\theta) : \theta \in \bigcup_{T\geq 1} \Theta_T\}$ is identifiable;
(4) The set of log-likelihood ratio tests $\mathcal{L}_T = \{\lambda(\theta_d^*, \theta_d; y) : \theta_d \in \Theta_T\}$, $\forall d \in \mathcal{D}$, satisfies

$$\lim_{T\to\infty} \sup_{f\in\mathcal{L}_T} \left| \frac{1}{T} \sum_{t=1}^{T} f(Y^{(t)}) - \mathbb{E}[f(Y)] \right| = 0, \quad a.s., \tag{EC.17}$$

where $Y \sim p(\cdot|\theta_d^*)$ and $Y^{(1)}, \ldots, Y^{(T)}$ are independent and identically distributed samples of Y. Moreover, we assume that

(5) For any $d \in \mathcal{D}$ and y in the support of $p(\cdot|\theta_d^*)$,

$$\lim_{T\to\infty} \frac{1}{T} \sup_{\theta_d\in\Theta_T} |\lambda(\theta_d^*, \theta_d; y)| = \lim_{T\to\infty} \frac{1}{T} \sup_{\theta_d\in\Theta_T} D(\theta_d^*\|\theta_d) = 0;$$

(6) For any $d \in \mathcal{D}$ and $\upsilon \geq 1$, the sequence of stopping times $\tau_{d,\upsilon} := \min\{\tau : \sum_{t=1}^{\tau} \psi_t(d)(1-\psi_t(d)) > \upsilon\}$ satisfy that $\tau_{d,\upsilon} < \infty$ and $\lim_{\upsilon\to\infty} \sqrt{\upsilon}/\tau_{d,\upsilon} = 0$, a.s.;
(7) For any $d \in \mathcal{D}$, we have

$$\lim_{T\to\infty} \frac{1}{T} \ln\left(\sup_{\theta_d\in\Theta_T} \|\partial D(\theta_d^*\|\theta_d)/\partial\theta_d\| \right) = 0.$$

Assumptions EC.1.(1-4) is a mimicry of Assumption 2 for increasing parameter spaces. Assumptions EC.1.(5,7) ensure that the likelihood ratio functions, the KL divergence functions and their derivatives would not diverge, which are easy to check for common parametric models. Assumption EC.1.(6) requires that the sampling ratio of any design is neither too large nor too small. It is harder to check this assumption theoretically than others. However, it can be checked numerically. Also, recall that TTTS-C asymptotically exhibits sampling ratios $\psi_t(d)$ at constant levels, implying that this assumption is realistic.

The following modification of the posterior large deviations principle is valid.

COROLLARY EC.1. *Under Assumptions 1 and EC.1, for any open set $\tilde{\mathbf{\Theta}} \subseteq \bigcup_{T\to\infty} \mathbf{\Theta}_T$, we have*

$$\lim_{T\to\infty} -\frac{1}{T}\ln \Pi_T(\tilde{\mathbf{\Theta}}\cap\mathbf{\Theta}_T) - \inf_{\boldsymbol{\theta}\in\tilde{\mathbf{\Theta}}\cap\mathbf{\Theta}_T} D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta}) = 0, \quad a.s.$$

Proof. Without loss of generality, assume that $\tilde{\mathbf{\Theta}} \subseteq \mathbf{\Theta}_1 \subseteq \mathbf{\Theta}_T$, $\forall T \geq 1$. We argue by repeating the proof of (EC.1) by showing (EC.5-EC.7), and repeating the proof of (EC.10). It can be easily seen that (EC.5) and (EC.7) still hold true with Θ replaced by Θ_T under Assumption EC.1.(4) and Assumption EC.1.(5), respectively.

The validity of (EC.6) with Θ replaced by Θ_T can be justified utilizing the martingale central limit theorem [21], which states that as $\upsilon\to\infty$,

$$\left|\frac{1}{\sqrt{\upsilon}}\sum_{t=1}^{\tau_\upsilon}(\xi_t(d)-\psi_t(d))\right| \xrightarrow{d} N(0,1),$$

where $\xrightarrow{d}$ denotes the weak convergence of distributions. The martingale central limit theorem holds true since $|\xi_t(d)-\psi_t(d)|$ is bounded. By assumption, we have $\sqrt{\upsilon}/\tau_\upsilon \to 0$, *a.s.*. Hence, the modification of (EC.6) follows from the Slutsky's lemma by putting pieces together.

Finally, it suffices to show (EC.10). Again, the left-hand side is non-negative by nature. For the other direction, denote $r_T := \sup_{\theta\in\Theta_T}\|\partial D(\theta^*\|\theta)/\partial\theta\|$. Choose $\varepsilon_T := \frac{k+1}{T\cdot r_T}$ and a finite ε_T-cover O of $\tilde{\Theta}$ such that each $o\in O$ intersects $\tilde{\Theta}$. The same inequality thus follows as

$$-\frac{1}{T}\ln\int_{\tilde{\mathbf{\Theta}}}\exp\{-TD_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})\}d\boldsymbol{\theta} - \inf_{\boldsymbol{\theta}\in\tilde{\mathbf{\Theta}}} D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta}) \leq -\frac{1}{T}\ln\int_{\tilde{\mathbf{\Theta}}\bigcap o_T} d\boldsymbol{\theta} + r_T\varepsilon_T,$$

where o_T is an ε_T-ball in O that contains $\boldsymbol{\theta}_T := \arg\min_{\boldsymbol{\theta}\in\mathbf{\Theta}_T} D_{\bar{\psi}_T}(\boldsymbol{\theta}^*\|\boldsymbol{\theta})$ in its closure. Let V_{k+1} denote the volume of the unit ball in $\mathbb{R}^{k+1}$. Therefore, it follows from Assumption EC.1.(7) that

$$-\frac{1}{T}\ln\int_{\tilde{\mathbf{\Theta}}\bigcap o_T} d\boldsymbol{\theta} + r_T\varepsilon_T = -\frac{1}{T}\ln V_{k+1} - \frac{k+1}{T}\ln\varepsilon_T + r_T\varepsilon_T = -\frac{1}{T}\ln V_{k+1} + \frac{k+1}{T}\ln\frac{eT\cdot r_T}{k+1} \to 0,$$

which justifies (EC.10) and thus completes the proof. $\square$

Appendix B: Consistency proof of TTTS-C. Then we establish Lemma EC.1, which intermediates between the posterior large deviations principle and the proof of consistency.

LEMMA EC.1. *Under Assumptions 1 and 2 (or 2′), with probability 1, if $\Psi_T(d) := T\bar{\psi}_T(d)\to\infty$ for some $d\in\mathcal{D}$, then for all $\varepsilon>0$ small enough, we have*

$$\Pi_T\left(\left\{\boldsymbol{\theta}\in\mathbf{\Theta}\middle| |\mu_d-\mu_d^*|\geq\varepsilon\right\}\right)\to 0, \tag{EC.18}$$

where $|\cdot|$ is the Euclidean norm; if $\lim_{T\to\infty}\Psi_T(d)<\infty$ for some $d\in\mathcal{D}$, then for any open set $(\mu',\mu'')\subseteq M$,

$$\inf_T\ \Pi_T\left(\{\boldsymbol{\theta}\in\mathbf{\Theta}|\mu_d\in(\mu',\mu'')\}\right)>0. \tag{EC.19}$$

Proof. Without loss of generality, assume $\boldsymbol{\theta}^*$ is an interior point of $\mathbf{\Theta}$. Now, fix a context $c\in\mathcal{C}$ and a design $d\in\mathcal{D}_c$. We will drop the subscripts d if no confusion rises. Recall that we assume an uninformative prior so that the marginal density of the posterior distribution in $\theta=\theta_d$ is

$$\Pi_{T,d}(\theta) = \frac{L_{T,d}(\theta)}{\int_\Theta L_{T,d}(\theta')d\theta'} = \frac{\exp\{-\sum_{t=1}^T\xi_t(d)\lambda(\theta^*,\theta;Y^{(t)})\}}{\int_\Theta\exp\{-\sum_{t=1}^T\xi_t(d)\lambda(\theta^*,\theta';Y^{(t)})\}d\theta'}. \tag{EC.20}$$

First, we consider the case when $\Psi_T(d) \to \infty$. Recall that Equation (EC.5) implies that with probability 1, if $\Psi_T(d) \to \infty$, then

$$\sup_{\theta \in \Theta} \left(\sum_{t=1}^{T} \xi_t(d) \right)^{-1} \left| \sum_{t=1}^{T} \xi_t(d) \lambda(\theta^*, \theta; Y^{(t)}) - \Psi_T(d) D(\theta^* \| \theta) \right| \to 0.$$

Together with Levy's extension of the Borel-Cantelli Lemma [55], which implies that $\sum_{t=1}^{T} \xi_t(d)/\Psi_T(d) \to 1$, we can conclude that, with probability 1, if $\Psi_T(d) \to \infty$, then

$$\sup_{\theta_d \in \Theta} \Psi_T^{-1}(d) \left| \sum_{t=1}^{T} \xi_t(d) \lambda(\theta^*, \theta; Y^{(t)}) - \Psi_T(d) D(\theta^* \| \theta) \right| \to 0.$$

Let v_T denote the supremum on the left-hand side, and then (EC.20) can be bounded by

$$\Pi_{T,d}(\theta) \le \frac{\exp\{-\Psi_T(d)(D(\theta^* \| \theta) - v_T)\}}{\int_{\Theta} \exp\{-\Psi_T(d)(D(\theta^* \| \theta') + v_T)\} d\theta'}. \tag{EC.21}$$

To bound the posterior probability in (EC.18), we investigate the behavior of the KL divergence. We will need a lemma of van der Vaart [54], which is restated as follows.

Lemma EC.2 **(Lemma 5.35 of [54]).** *Let $\theta, \theta' \in \Theta$. If $p(\cdot|\theta)$ differs from $p(\cdot|\theta')$ with positive positive Lebesgue measure, then $D(\theta \| \theta') > 0$.*

Let $F := \inf_{\theta \in \Theta: |\mu - \mu^*| \ge \varepsilon} D(\theta^* \| \theta)$. For each $\theta = (\mu, \eta)$ such that $|\mu - \mu^*| \ge \varepsilon$, it follows from either Assumption 2.(3) or 2′.(3′) that $p(\cdot|\theta^*)$ differs from $p(\cdot|\theta)$ on a set with positive Lebesgue measure. According to Lemma EC.2, we have $D(\theta^* \| \theta) > D(\theta^* \| \theta^*) = 0$. Since every continuous function on a compact set attains its infimum, we conclude that $F > 0$. Recall that $r = \sup_{\theta \in \Theta} \| \partial D(\theta^* \| \theta)/\partial \theta \| \le \infty$. Then for $\varepsilon' = F/2r$, and for any θ such that $\|\theta - \theta^*\|_2 \le \varepsilon'$,

$$D(\theta^* \| \theta) \le D(\theta^* \| \theta^*) + r \|\theta - \theta^*\| \le F/2.$$

According to (EC.21), we see that

$$\begin{aligned} \Pi_{T,d}\left(\{ \boldsymbol{\theta} \in \boldsymbol{\Theta} \big| |\mu - \mu^*| \ge \varepsilon \} \right) &\le \frac{\int_{\{|\mu - \mu^*| \ge \varepsilon\}} \exp\{-\Psi_T(d)(D(\theta^* \| \theta) - v_T)\} d\theta}{\int_{\{\|\theta - \theta^*\|_2 \le \varepsilon'\}} \exp\{-\Psi_T(d)(D(\theta^* \| \theta) + v_T)\} d\theta} \\ &\le \frac{\mathrm{Diam}(\Theta)^{d+1} \exp\left\{-\Psi_T(d)\big(F - v_T\big)\right\}}{(\varepsilon')^{d+1} \exp\left\{-\Psi_T(d)\big(F/2 + v_T\big)\right\}}, \end{aligned}$$

where $\mathrm{Diam}(\cdot)$ indicates the diameter of sets. The bound simplifies to $(2r\,\mathrm{Diam}(\Theta)/F)^{d+1} \times \exp\{-\Psi_T(d)(F/2 - 2v_T)\}$, which vanishes almost surely since $\Psi_T \to \infty$, $v_T \to 0$ and $F > 0$.

Now consider the second part. Again by Levy's extension of the Borel-Cantelli Lemma, with probability 1, if $\lim_{T\to\infty} \Psi_T(d) < \infty$, then $\lim_{T\to\infty} \sum_{t=1}^{T} \xi_t(d) < \infty$. Fix any realization of $\xi_t(d)$'s and $Y^{(t)}$'s with $\lim_{T\to\infty} \sum_{t=1}^{T} \xi_t(d) < \infty$, we have $\sup_{\theta \in \Theta} \sup_{T \in \mathbb{N}} \left| \sum_{t=1}^{T} \xi_t(d) \lambda(\theta^*, \theta; Y^{(t)}) \right| < \infty$ since the summation only involves a finite number of terms, λ is continuous in θ and Θ is compact. By (EC.20), $\delta < \Pi_{T,d}(\theta) < \delta^{-1}$ is bounded both from below and above for some $\delta > 0$, and (EC.19) follows immediately. □

Lemma EC.1 shows that the posterior distribution concentrates around the true value as long as the cumulative sampling ratio put into the design is not finite ultimately. The converse is also true as in (EC.19), which is the key to the consistency result. Notice that this result is not a trivial corollary of Theorem 1 since we do not require a positive ratio of samples. Actually, the identifiable condition in Assumption 2 or 2′ is a necessity in establishing (EC.18).

Then we can complete the proof for Theorem 3 by showing a stronger result as follows.

Theorem EC.1. *Suppose Assumptions 1 and 2 (or 2′) hold. Suppose tuning hyperparameter $\gamma_t(c)$ is adaptive to $\{\mathcal{E}_t\}$, and then the proposed TTTS-C with tuning hyperparameter $\gamma_t(c)$ is consistent, i.e., $\sum_{t=1}^T \xi_t(d) \to \infty$, a.s., as $T \to \infty$, $\forall c \in \mathcal{C}$, $\forall d \in \mathcal{D}_c$.*

Proof. We argue by contradiction. To this end, suppose the set of inadequately sampled context-design pairs $\mathscr{I} = \{(c,d) : c \in \mathcal{C}, d \in \mathcal{D}_c, \lim_{T\to\infty} \sum_{t=1}^T \xi_t(d) < \infty\}$ is non-empty. We first assume that $\sum_{t=1}^\infty \gamma_t(c) = \infty$. Let $\hat{\boldsymbol{\theta}}_{2,k}^{(t)}$, $\forall k \geq 1$ denote the independent copies of $\hat{\boldsymbol{\theta}}_2^{(t)}$ generated in Algorithm 1 or Algorithm 2. The probability of sampling on context c and design d conditioned on $\mathcal{E}_t$ at time t can be lower bounded by

$$\mathbb{P}(\xi_t(d)|\mathcal{E}_t) \geq \mathbb{P}\left(\left(\hat{\mu}_1^{(t)}\right)_d > \max_{d' \in \mathcal{D}_c \setminus \{d\}} \left(\hat{\mu}_1^{(t)}\right)_{d'} \middle| \mathcal{E}_t\right) \cdot \mathbb{P}\left(\left(\hat{\mu}_{2,1}^{(t)}\right)_d < \min_{d' \in \mathcal{D}_c \setminus \{d\}} \left(\hat{\mu}_{2,1}^{(t)}\right)_{d'} \middle| \mathcal{E}_t\right) \cdot \gamma_t(c) \frac{1}{|\mathcal{D}_c|} \cdot \frac{1}{|\mathcal{C}|}. \tag{EC.22}$$

This inequality holds because the right-hand side provides a lower bound for the probability of the event where the posterior sample $\hat{\boldsymbol{\theta}}_1^{(t)}$ indicates that design d is the best design in context c, and the first independent copy $\hat{\boldsymbol{\theta}}_{2,1}^{(t)}$ of the posterior sample $\hat{\boldsymbol{\theta}}_2^{(t)}$ indicates that design d is the worst design in context c, so that context c and design d will be simulated. The factor $1/|\mathcal{C}|$ arises since Δ_t may also include contexts other than c in the event we described above, and similarly, for the factor $1/|\mathcal{D}_c|$.

By Lemma EC.1, for any $(c,d) \in \mathcal{I}$, and some sufficiently small number $\varepsilon > 0$, there exists $\delta > 0$ and $\bar{T} \geq 1$ such that for any $t \geq \bar{T}$,

$$\mathbb{P}\left(\left(\hat{\mu}_1^{(t)}\right)_d > \max_{d' \in \mathcal{D}_c \setminus \{d\}} \left(\hat{\mu}_1^{(t)}\right)_{d'} \middle| \mathcal{E}_t\right) \geq \Pi_t\left(\bigcap_{d' \neq d} \left\{\mu_{d'} < \mu^*_{d^*(c)} + \varepsilon/2\right\} \bigcap \left\{\mu_d > \mu^*_{d^*(c)} + \varepsilon\right\}\right) > \delta,$$

and similarly,

$$\mathbb{P}\left(\left(\hat{\mu}_{2,1}^{(t)}\right)_d < \min_{d' \in \mathcal{D}_c \setminus \{d\}} \left(\hat{\mu}_{2,1}^{(t)}\right)_{d'} \middle| \mathcal{E}_t\right) \geq \Pi_t\left(\bigcap_{d' \neq d} \left\{\mu_{d'} > \mu^*_{d^*(c)} - \varepsilon/2\right\} \bigcap \left\{\mu_d < \mu^*_{d^*(c)} - \varepsilon\right\}\right) > \delta.$$

Then we see that

$$\sum_{t=\bar{T}}^\infty \psi_t(d) = \sum_{t=\bar{T}}^\infty \mathbb{P}(\xi_t(d)|\mathcal{E}_t) \geq \frac{\delta^2}{|\mathcal{C}||\mathcal{D}|} \sum_{t=\bar{T}}^\infty \gamma_t(c)$$

diverges according to the assumption $\sum_{t=1}^\infty \gamma_t(c) = \infty$. Conversely, if $\sum_{t=1}^\infty \gamma_t(c) < \infty$, then $\sum_{t=1}^\infty 1 - \gamma_t(c) = \infty$. It follows from the same argument that

$$\mathbb{P}(\xi_t(d)|\mathcal{E}_t) \geq \mathbb{P}\left(\left(\hat{\mu}_1^{(t)}\right)_d < \max_{d' \in \mathcal{D}_c \setminus \{d\}} \left(\hat{\mu}_1^{(t)}\right)_{d'} \middle| \mathcal{E}_t\right) \cdot \mathbb{P}\left(\left(\hat{\mu}_{2,1}^{(t)}\right)_d > \max_{d' \in \mathcal{D}_c \setminus \{d\}} \left(\hat{\mu}_{2,1}^{(t)}\right)_{d'} \middle| \mathcal{E}_t\right) \cdot \frac{1 - \gamma_t(c)}{|\mathcal{D}_c| \cdot |\mathcal{C}|}. \tag{EC.23}$$

Similarly, there exists $\delta > 0$ such that the latter two probabilities in (EC.23) are larger than or equal to δ for $t \geq \bar{T}$, and therefore,

$$\sum_{t=\bar{T}}^\infty \psi_t(d) = \sum_{t=\bar{T}}^\infty \mathbb{P}(\xi_t(d)|\mathcal{E}_t) \geq \frac{\delta^2}{|\mathcal{C}||\mathcal{D}|} \sum_{t=\bar{T}}^\infty (1 - \gamma_t(c))$$

diverges as well.

By Levy's extension of the Borel-Cantelli lemma [55], $\lim_{T\to\infty} \sum_{t=1}^T \xi_t(d) < \infty$ and $\lim_{T\to\infty} \sum_{t=1}^T \psi_t(d) = \infty$ holds simultaneously with zero probability. However, $\lim_{T\to\infty} \sum_{t=1}^T \xi_t(d) < \infty$ implies $\lim_{T\to\infty} \sum_{t=1}^T \psi_t(d) = \infty$ for any sample path. Hence the probability of the even $\lim_{T\to\infty} \sum_{t=1}^T \xi_t(d) < \infty$ must be zero, which completes the proof. □

Appendix C: Characterizing the optimal sampling ratios.

Proof of Proposition 1. (a) Fix the second argument y for $G_d(x,y)$. By definition, $G_d(x,y) = \inf_{(\theta_d,\theta_{d^*(c)}):\mu_d\geq\mu_{d^*(c)}} xD(\theta_d^*\|\theta_d) + yD(\theta^*_{d^*(c)}\|\theta_{d^*(c)})$. Suppose for $\tilde{x}$ and $\breve{x}$ with $\tilde{x} < \breve{x}$, the infimum is taken at $(\tilde{\theta}_d, \tilde{\theta}_{d^*(c)})$ and $(\breve{\theta}_d, \breve{\theta}_{d^*(c)})$, respectively. Then we see that

$$\begin{aligned} G_d(\tilde{x},y) &= \tilde{x}D(\theta_d^*\|\tilde{\theta}_d) + yD(\theta^*_{d^*(c)}\|\tilde{\theta}_{d^*(c)}) \\ &\leq \tilde{x}D(\theta_d^*\|\breve{\theta}_d) + yD(\theta^*_{d^*(c)}\|\breve{\theta}_{d^*(c)}) \\ &\leq \breve{x}D(\theta_d^*\|\breve{\theta}_d) + yD(\theta^*_{d^*(c)}\|\breve{\theta}_{d^*(c)}) \\ &= G_d(\breve{x},y), \end{aligned}$$

where the first inequality follows from the definition of $(\tilde{\theta}_d, \tilde{\theta}_{d^*(c)})$, and the second follows from the non-negativity of KL divergence. For the second argument, a symmetric proof applies.

(b) The first statement follows immediately by observing that $G_d(x,y)$ is the infimum of a set of linear functions in (x,y). Then the local continuity at any inner point $(x,y)\in(\mathbb{R}^+)^2$ follows consequently. Now it suffices to show that the local continuity at the boundary of $(\mathbb{R}^+)^2$ holds. Without loss of generality, suppose $\{(x_n,y_n)\}_{n\geq1}$ is a sequence of inner points of $(\mathbb{R}^+)^2$, i.e., $x_n > 0$ and $y_n > 0$, with $x_n \to x_0 \geq 0$ and $y_n \to 0$. By $G_d(x_n,y_n) \leq x_nD(\theta_d^*\|\theta_d^*) + y_nD(\theta^*_{d^*(c)}\|\theta_d^*) \leq y_n \sup_{\theta\in\Theta} D(\theta^*_{d^*(c)}\|\theta) \to 0$, we see that $G_d(x_n,y_n) \to G_d(x_0,0) = 0$. To sum up, $G_d(\cdot,\cdot)$ is a continuous function.

(c) It follows immediately from the definition

$$\begin{aligned} G_d(hx,hy) &= \inf_{(\theta_d,\theta_{d^*(c)}):\mu_d\geq\mu_{d^*(c)}} hxD(\theta_d^*\|\theta_d) + hyD(\theta^*_{d^*(c)}\|\theta_{d^*(c)}) \\ &= h\inf_{(\theta_d,\theta_{d^*(c)}):\mu_d\geq\mu_{d^*(c)}} xD(\theta_d^*\|\theta_d) + yD(\theta^*_{d^*(c)}\|\theta_{d^*(c)}) \\ &= hG_d(x,y). \quad \square \end{aligned}$$

Proof of Proposition 2. The optimization problem (8) can be rewritten as

$$\begin{aligned} \Gamma^* = \max_{\alpha,\beta\geq0} \quad & z \\ s.t. \quad & z \leq \alpha(c)z_c, && \forall c\in\mathcal{C}, \\ & z_c \leq G_d\left(\beta(c,d^*(c)),\beta(c,d)\right), && \forall c\in\mathcal{C}, \forall d\in\mathcal{D}_c\backslash\{d^*(c)\}, \\ & \sum_{c\in\mathcal{C}}\alpha(c)=1, \quad \sum_{d\in\mathcal{D}_c}\beta(c,d)=1, && \forall c\in\mathcal{C}. \end{aligned} \tag{EC.24}$$

Therein, decision variables z and z_c, $c\in\mathcal{C}$ are introduced to facilitate analysis. The KKT conditions provide sufficient and necessary conditions [44] for (EC.24), since the constraints are apparently strictly feasible and satisfy the Slater's condition. The KKT conditions are as follows

$$\begin{gathered} \sum_{d'\in\mathcal{D}_c\backslash\{d^*(c)\}} \Lambda_2(c,d')\frac{\partial}{\partial x_1}G_{d'}(\beta(c,d^*(c)),\beta(c,d')) - \Lambda_4(c) = 0, \quad \forall c\in\mathcal{C}, \\ \Lambda_2(c,d)\frac{\partial}{\partial x_2}G_d(\beta(c,d^*(c)),\beta(c,d)) - \Lambda_4(c) = 0, \quad \forall c\in\mathcal{C}, \forall d\in\mathcal{D}_c\backslash\{d^*(c)\}, \\ 1-\sum_{c\in\mathcal{C}}\Lambda_1(c)=0, \quad \Lambda_1(c)z_c - \Lambda_3 = 0, \quad \Lambda_1(c)\alpha(c) - \sum_{d'\in\mathcal{D}_c\backslash\{d^*(c)\}}\Lambda_2(c,d') = 0, \quad \forall c\in\mathcal{C}, \\ \Lambda_1(c)\left(\alpha(c)z_c - z\right) = 0, \quad \Lambda_2(c,d)\left(G_d(\beta(c,d^*(c)),\beta(c,d)) - z_c\right) = 0, \quad \forall c\in\mathcal{C}, \forall d\in\mathcal{D}_c\backslash\{d^*(c)\}, \end{gathered}$$

where $\Lambda_1(c)\geq0, \Lambda_2(c,d)\geq0, \Lambda_3\in\mathbb{R}$, and $\Lambda_4(c)\in\mathbb{R}$ are dual variables, and for all $c\in\mathcal{C}$ and $d\in\mathcal{D}_c$,

$$\left(-\frac{\partial}{\partial x_1}G_d(\beta(c,d^*(c)),\beta(c,d)), -\frac{\partial}{\partial x_2}G_d(\beta(c,d^*(c)),\beta(c,d))\right)$$

denotes a subgradient of $-G_d(\cdot,\cdot)$ evaluated at $(\beta(c,d^*(c)),\beta(c,d))$.

Note that any optimal solution α^*,β^* together with z^*,z_c^* is strictly positive. According to $\Lambda_1(c)z_c^*+\Lambda_3=0$, we have either $\Lambda_1(c)=\Lambda_3=0$, $\forall c\in\mathcal{C}$, or $\Lambda_1(c)>0$ and $\Lambda_3\neq 0$, $\forall c\in\mathcal{C}$. However, the former contradicts $1-\sum_{c'\in\mathcal{C}}\Lambda_1(c')=0$, hence $\Lambda_1(c),\Lambda_3\neq 0$.

Finally, we show that $\Lambda_2(c,d)>0$ and $\Lambda_4(c)\neq 0$, $\forall c\in\mathcal{C}, d\in\mathcal{D}_c$. Again assume for the sake of contradiction that $\Lambda_4(c)=0$ for some $c\in\mathcal{C}$. It follows from the monotonicity of $G_d(\cdot,\cdot)$ in both arguments that $\partial G_d/\partial x_i\geq 0$, $i=1,2$. Then following the first two conditions and the non-negativity of Λ_2 and $\partial G_d/\partial x_i$, $i=1,2$, we have

$$\Lambda_2(c,d)\frac{\partial}{\partial x_1}G_d(\beta^*(c,d^*(c)),\beta^*(c,d))=\Lambda_2(c,d)\frac{\partial}{\partial x_2}G_d(\beta^*(c,d^*(c)),\beta^*(c,d))=0.$$

Note that $(0,0)$ is not a subgradient of $-G_d(\cdot,\cdot)$ due to the fact that $G_d(\cdot,\cdot)$ is homogeneous of degree 1. Therefore, $\Lambda_2(c,d)=0$, $\forall d\in\mathcal{D}_c$, leading to $0=\sum_{d'\in\mathcal{D}_c\setminus\{d^*(c)\}}\Lambda_2(c,d')=\Lambda_1(c)\alpha^*(c)>0$. It turns out that $\Lambda_2(c,d)>0$ and $\Lambda_4(c)\neq 0$, $\forall c\in\mathcal{C}, d\in\mathcal{D}_c$.

According to the second line of the KKT conditions, $\Lambda_4(c)\neq 0$ implies $\partial G_d(\beta^*(c,d^*(c)),\beta^*(c,d))/\partial x_2\neq 0$, $\forall c\in\mathcal{C},d\in\mathcal{D}_c$. Combining the first two KKT conditions, we have

$$\sum_{d\in\mathcal{D}_c\setminus\{d^*(c)\}}\frac{\partial G_d(\beta^*(c,d^*(c)),\beta^*(c,d))/\partial x_1}{\partial G_d(\beta^*(c,d^*(c)),\beta^*(c,d))/\partial x_2}=1,\quad \forall c\in\mathcal{C}.$$

Moreover, $\Lambda_1(c)>0$, $\Lambda_2(c,d)>0$, and the last line of the KKT conditions yield

$$\alpha^*(c)G_d(\beta^*(c,d^*(c)),\beta^*(c,d))=z^*,\quad \forall c\in\mathcal{C},d\in\mathcal{D}_c\setminus\{d^*(c)\}.$$

On the other hand, we show that any feasible solution (α,β,z) satisfying the conditions in Proposition 2 constitutes an optimal solution to (EC.24). Actually, it follows simply by choosing $z_c=z/\alpha(c)$, $\Lambda_1(c)=\alpha(c)$, $\Lambda_2(c,d)=\Lambda_4(c)/(\partial G_d(\beta(c,d^*(c)),\beta(c,d))/\partial x_2)$, $\Lambda_3=z$, and $\Lambda_4(c)=(\alpha(c))^2/\sum_{d'\in\mathcal{D}_c\setminus\{d^*(c)\}}(\partial G_d(\beta(c,d^*(c)),\beta(c,d'))/\partial x_2)^{-1}$. □

Similar arguments work for the following two propositions.

Proof of Proposition 3. The optimization problem (11) can be written similarly as (EC.24) with an additional constraint $\beta(c,d^*(c))=\gamma(c)$.

$$\begin{aligned}
\Gamma^*=\max_{\alpha,\beta\geq 0}\quad & z && \\
s.t.\quad & z\leq\alpha(c)z_c, && \forall c\in\mathcal{C},\\
& z_c\leq G_d\left(\beta(c,d^*(c)),\beta(c,d)\right), && \forall c\in\mathcal{C},\forall d\in\mathcal{D}_c\setminus\{d^*(c)\},\\
& \sum_{c\in\mathcal{C}}\alpha(c)=1,\quad \sum_{d\in\mathcal{D}_c}\beta(c,d)=1, && \forall c\in\mathcal{C},\\
& \beta(c,d^*(c))=\gamma(c), && \forall c\in\mathcal{C}.
\end{aligned}\tag{EC.25}$$

Then the corresponding KKT conditions coincide with those of (EC.24), with the only exception that the first line of the KKT conditions for (EC.24) is replaced by feasible conditions $\beta(c,d^*(c))=\gamma(c)$. The corresponding KKT conditions are as follows:

$$\begin{gathered}
\beta(c,d^*(c))=\gamma(c),\quad \forall c\in\mathcal{C},\\
\Lambda_2(c,d)\frac{\partial}{\partial x_2}G_d(\beta(c,d^*(c)),\beta(c,d))-\Lambda_4(c)=0,\quad \forall c\in\mathcal{C},\forall d\in\mathcal{D}_c\setminus\{d^*(c)\},\\
1-\sum_{c\in\mathcal{C}}\Lambda_1(c)=0,\quad \Lambda_1(c)z_c-\Lambda_3=0,\quad \Lambda_1(c)\alpha(c)-\sum_{d'\in\mathcal{D}_c\setminus\{d^*(c)\}}\Lambda_2(c,d')=0,\quad \forall c\in\mathcal{C},\\
\Lambda_1(c)\left(\alpha(c)z_c-z\right)=0,\quad \Lambda_2(c,d)\left(G_d(\beta(c,d^*(c)),\beta(c,d))-z_c\right)=0,\quad \forall c\in\mathcal{C},\forall d\in\mathcal{D}_c\setminus\{d^*(c)\}.
\end{gathered}$$

For any feasible solution (α, β, z) satisfying (12), the KKT conditions are met automatically by choosing $z_c = z/\alpha(c)$, $\Lambda_1 = \alpha(c)$, $\Lambda_2(c,d) = \Lambda_4(c)/(\partial G_d(\beta(c,d^*(c)), \beta(c,d))/\partial x_2)$, $\Lambda_3 = z$, and $\Lambda_4(c) = (\alpha(c))^2/\sum_{d' \in \mathcal{D}_c \setminus \{d^*(c)\}} (\partial G_d(\beta(c,d^*(c)), \beta(c,d'))/\partial x_2)^{-1}$.

On the other hand, consider an optimal solution $(\alpha^*, \beta^*, z_c^*, z^*)$ to (EC.25). We will construct an optimal solution that satisfies (12). Using the same argument in the proof of Proposition 2, we see that $\Lambda_1(c) > 0, \Lambda_2(c,d) > 0$, $\forall c \in \mathcal{C}, d \in \mathcal{D}_c$. It follows from the last line of the KKT conditions that $\alpha^*(c) = (z_c^*)^{-1}/\sum_{c' \in \mathcal{C}} (z_{c'}^*)^{-1}$ and $z^* = 1/\sum_{c' \in \mathcal{C}} (z_{c'}^*)^{-1}$.

We claim that for each $c \in \mathcal{C}$, there exists $\tilde{\beta}$ such that $G_d(\gamma(c), \tilde{\beta}(c,d)) = z_c^*$ for all $d \in \mathcal{D}_c \setminus \{d^*(c)\}$. This is trivial if $G_d(\gamma(c), \beta^*(c,d)) = z_c^*$ for all $d \in \mathcal{D}_c \setminus \{d^*(c)\}$. On the contrary, if $G_d(\gamma(c), \beta^*(c,d)) > z_c^*$ for some $d \in \mathcal{D}_c$ and $c \in \mathcal{C}$, then for this fixed c and d, there exists $d' \in \mathcal{D}_c \setminus \{d^*(c), d\}$ such that $G_{d'}(\gamma(c), \beta) = z_c^*$ for all $\beta \geq \beta^*(c,d')$. Otherwise, for each $d' \in \mathcal{D}_c \setminus \{d^*(c), d\}$, there exists $\beta_0 > \beta^*(c,d')$ such that $G_{d'}(\gamma(c), \beta_0) > G_{d'}(\gamma(c), \beta^*(c,d'))$. According to Proposition 1, $G_{d'}(\cdot,\cdot)$ is concave, and thus for $\varepsilon \leq \beta_0 - \beta^*(c,d')$

$$\frac{G_{d'}(\gamma(c), \beta^*(c,d') + \varepsilon) - G_{d'}(\gamma(c), \beta^*(c,d'))}{\varepsilon} \geq \frac{G_{d'}(\gamma(c), \beta_0) - G_{d'}(\gamma(c), \beta^*(c,d'))}{\beta_0 - \beta^*(c,d')} > 0.$$

The solution $\bar{\beta}(c,d) = \beta^*(c,d) - (|\mathcal{D}_c| - 2)\varepsilon > 0$ and $\bar{\beta}(c,d') = \beta^*(c,d') + \varepsilon > 0$ satisfies that for $d' \in \mathcal{D}_c \setminus \{d^*(c), d\}$,

$$G_{d'}(\gamma(c), \bar{\beta}(c,d')) > G_{d'}(\gamma(c), \beta^*(c,d')) \geq z_c^*.$$

Since $G_d(\cdot,\cdot)$ is uniformly continuous, $G_d(\gamma(c), \beta^*(c,d)) > z_c^*$ and $|\bar{\beta}(c,d) - \beta^*(c,d)| \leq (|\mathcal{D}_c| - 2)\varepsilon$, we can choose $\varepsilon > 0$ to be small enough so that $G_d(\gamma(c), \bar{\beta}(c,d)) > z_c^*$. Therefore,

$$\min_{d \in \mathcal{D}_c \setminus \{d^*(c)\}} G_d(\gamma(c), \bar{\beta}(c,d)) > \min_{d \in \mathcal{D}_c \setminus \{d^*(c)\}} G_d(\gamma(c), \beta^*(c,d)) = z_c^*.$$

Hence, let $\bar{z}_c := \min_{d \in \mathcal{D}_c \setminus \{d^*(c)\}} G_d(\gamma(c), \bar{\beta}(c,d))$, $\bar{z}_{c'} := z_{c'}^*$, and $\forall c' \in \mathcal{C} \setminus \{c\}$, and define $\bar{\alpha}(c') := \bar{z}_{c'}^{-1}/\sum_{c'' \in \mathcal{C}} \bar{z}_{c''}^{-1}$, $\forall c' \in \mathcal{C}$ and $\bar{z} := 1/\sum_{c'' \in \mathcal{C}} \bar{z}_{c''}^{-1}$. Then $\bar{z} > z^*$, and thus $(\bar{\alpha}, \bar{\beta}, \bar{z}_c, \bar{z})$ yields a feasible solution to (EC.25) with a larger objective than z^*, which contradicts the optimality.

Then, we fix $d' \in \mathcal{D}_c \setminus \{d^*(c), d\}$ with $G_{d'}(\gamma(c), \beta) = z_c^*$ for all $\beta \geq \beta^*(c,d')$. Finally, we define $\tilde{\beta}(c,d)$ implicitly by an arbitrary solution to the equation $G_d(\gamma(c), \tilde{\beta}(c,d)) = z_c^*$ for $d \in \mathcal{D}_c \setminus \{d^*(c), d'\}$. Note that it's well-defined by the intermediate value theorem since $G_d(\gamma(c), 0) = 0 < z_c^*$, $G_d(\gamma(c), \beta^*(c,d)) \geq z_c^*$, and $G_d(\cdot,\cdot)$ is continuous. Define $\tilde{\beta}(c,d') = 1 - \gamma(c) - \sum_{d \in \mathcal{D}_c \setminus \{d^*(c), d'\}} \tilde{\beta}(c,d)$, and then $(\alpha^*, \tilde{\beta}, z_c^*, z^*)$ is an optimal solution to (11) that satisfies the optimality condition (12).

The last statement follows immediately by noting that when $\gamma(c) = \beta^*(c, d^*(c))$, where (α^*, β^*) constitutes an optimal solution to (8), then an optimal solution of (11) is also optimal to (8). The conclusion follows immediately from Proposition 2. $\square$

Proof of Proposition 4. The optimization problem (18) can be rewritten as

$$\begin{aligned}
\Gamma^* = \max_{\alpha, \beta \geq 0} \quad & z \\
s.t. \quad & z \leq \alpha(c) z_c, && \forall c \in \mathcal{C}, \\
& z_c \leq G_{d,d'}\left(\beta(c,d), \beta(c,d')\right), && \forall c \in \mathcal{C}, \forall d \in \mathcal{P}_c, d \in \mathcal{U}_c, \\
& \sum_{c \in \mathcal{C}} \alpha(c) = 1, \quad \sum_{d \in \mathcal{D}_c} \beta(c,d) = 1, && \forall c \in \mathcal{C}.
\end{aligned} \tag{EC.26}$$

The first order conditions and complementary conditions with regard to the decision variable β and the second inequality constraint are given by

$$\sum_{d' \in \mathcal{U}_c} \Lambda_2(c,d,d') \frac{\partial}{\partial x_1} G_{d,d'}(\beta(c,d), \beta(c,d')) - \Lambda_4(c) = 0, \quad \forall c \in \mathcal{C}, \forall d \in \mathcal{P}_c, \tag{EC.27}$$

$$\sum_{d \in \mathcal{P}_c} \Lambda_2(c,d,d') \frac{\partial}{\partial x_2} G_{d,d'}(\beta(c,d), \beta(c,d')) - \Lambda_4(c) = 0, \quad \forall c \in \mathcal{C}, \forall d' \in \mathcal{U}_c, \tag{EC.28}$$

and

$$\Lambda_2(c,d,d')\left(G_{d,d'}(\beta(c,d),\beta(c,d'))-z_c\right)=0,\quad \forall c\in\mathcal{C},\forall d\in\mathcal{P}_c, d'\in\mathcal{U}_c, \tag{EC.29}$$

where $\partial/\partial x_i$'s are defined as in Proposition 2, and $\Lambda_2(c,d,d')$ and $\Lambda_4(c)$ are non-negative dual variables. Recall that the rate functions $G_{d,d'}(\cdot,\cdot)$ are homogeneous of degree 1, and thus for any $r>0$,

$$\begin{aligned}
&(r-1)G_{d,d'}(\beta(c,d),\beta(c,d'))\\
=&G_{d,d'}(r\beta(c,d),r\beta(c,d'))-G_{d,d'}(\beta(c,d),\beta(c,d'))\\
\le&(r-1)\beta(c,d)\frac{\partial}{\partial x_1}G_{d,d'}(\beta(c,d),\beta(c,d'))+(r-1)\beta(c,d')\frac{\partial}{\partial x_2}G_{d,d'}(\beta(c,d),\beta(c,d')),
\end{aligned}$$

which can only happen when

$$G_{d,d'}(\beta(c,d),\beta(c,d'))=\beta(c,d)\frac{\partial}{\partial x_1}G_{d,d'}(\beta(c,d),\beta(c,d'))+\beta(c,d')\frac{\partial}{\partial x_2}G_{d,d'}(\beta(c,d),\beta(c,d')).$$

Multiplying (EC.27) and (EC.28) with $\beta(c,d)$ and $\beta(c,d')$, respectively, and summing up these equations, we see that

$$\sum_{d\in\mathcal{P}_c}\sum_{d'\in\mathcal{U}_c}\Lambda_2(c,d,d')G_{d,d'}(\beta(c,d),\beta(c,d'))=\Lambda_4(c)\sum_{d\in\mathcal{P}_c}\beta(c,d)+\Lambda_4(c)\sum_{d'\in\mathcal{U}_c}\beta(c,d')=\Lambda_4(c).$$

According to (EC.29),

$$\sum_{d\in\mathcal{P}_c}\sum_{d'\in\mathcal{U}_c}\Lambda_2(c,d,d')z_c=\Lambda_4(c).$$

Note the fact that any optimal solution $(\alpha^*,\beta^*,z^*,z_c^*)$ to (EC.26) is positive, we see that either $\Lambda_4(c)>0$ or $\Lambda_2(c,d,d')=0$ for all $d\in\mathcal{P}_c$ and $d'\in\mathcal{U}_c$. The latter, however, is impossible due to the following KKT conditions

$$1-\sum_{c\in\mathcal{C}}\Lambda_1(c)=0,\quad \Lambda_1(c)z_c-\Lambda_3=0,\quad \Lambda_1(c)\alpha(c)-\sum_{d\in\mathcal{P}_c}\sum_{d'\in\mathcal{U}_c}\Lambda_2(c,d,d')=0,$$

where $\Lambda_1(c)$ and Λ_3 are non-negative dual variables.

Since $\Lambda_4(c)>0$, combining (EC.27) and (EC.28), for any $d\in\mathcal{P}_c$ and $d'\in\mathcal{U}_c$,

$$\max_{\tilde{d}\in\mathcal{P}_c}\Lambda_2(c,\tilde{d},d')>0,\quad \text{and}\quad \max_{\tilde{d}'\in\mathcal{U}_c}\Lambda_2(c,d,\tilde{d}')>0.$$

Together with the complementary slackness condition (EC.29), we have

$$\min_{\tilde{d}\in\mathcal{P}_c}G_{\tilde{d},d'}(\beta(c,\tilde{d}),\beta(c,d'))=\min_{\tilde{d}'\in\mathcal{U}_c}G_{d,\tilde{d}'}(\beta(c,d),\beta(c,\tilde{d}'))=z_c.$$

The remaining part of the proof is a simple tautology of the proof for Proposition 2, which we will omit to avoid redundancy. □

Appendix D: Rate-optimality of TTTS-C. In this section, we establish the asymptotic optimality of TTTS-C. Again, we first focus on the case of selecting contextual best designs where $m_c=1$. We prove Lemma 1, and Lemma 2 in Section D.1, and show Theorem 2 and Corollary 2 in Section D.2.

D.1. Sampling ratios of TTTS-C. The three lemmas below derive the (conditional) sampling ratio put into each design for TTTS-C. Through marginalizing $\hat{\boldsymbol{d}}_1^{(t)}$ and $\hat{\boldsymbol{d}}_2^{(t)}$ in (2), we obtain the sampling effort put into each design and context conditioned on $\mathcal{E}_{t-1}$ in Lemma EC.3.

Lemma EC.3. *Suppose Assumptions 1 and 2 (or 2′) hold. For any fixed $c \in \mathcal{C}$ and $d \in \mathcal{D}_c$, the sampling policy of TTTS-C at time t is*

$$\begin{aligned}\psi_t(d) = &\sum_{\boldsymbol{d}_1: d_1(c)=d} \prod_{c' \in \mathcal{C}} \pi_{d_1(c')}^{(t-1)} \sum_{c \in \Delta} \frac{\prod_{c' \in \Delta}\left(1-\pi_{d_1(c')}^{(t-1)}\right)\prod_{c' \notin \Delta} \pi_{d_1(c')}^{(t-1)}}{1-\prod_{c' \in \mathcal{C}} \pi_{d_1(c')}^{(t-1)}} \frac{\gamma(c)}{|\Delta|} \\ &+ \sum_{\boldsymbol{d}_1: d_1(c) \neq d} \prod_{c' \in \mathcal{C}} \pi_{d_1(c')}^{(t-1)} \sum_{c \in \Delta} \frac{\prod_{c' \in \Delta}\left(1-\pi_{d_1(c')}^{(t-1)}\right)\prod_{c' \notin \Delta} \pi_{d_1(c')}^{(t-1)}}{1-\prod_{c' \in \mathcal{C}} \pi_{d_1(c')}^{(t-1)}} \cdot \frac{\pi_d^{(t-1)}}{1-\pi_{d_1(c)}^{(t-1)}} \frac{1-\gamma(c)}{|\Delta|},\end{aligned} \tag{EC.30}$$

$$\alpha_t(c) = \sum_{\boldsymbol{d}_1} \frac{\prod_{c' \in \mathcal{C}} \pi_{d_1(c')}^{(t-1)}}{1-\prod_{c' \in \mathcal{C}} \pi_{d_1(c')}^{(t-1)}} \sum_{c \in \Delta} \prod_{c' \in \Delta}\left(1-\pi_{d_1(c')}^{(t-1)}\right) \prod_{c' \notin \Delta} \pi_{d_1(c')}^{(t-1)} \frac{1}{|\Delta|}, \tag{EC.31}$$

$$\beta_t(c,d) = \psi_t(d)/\alpha_t(c), \tag{EC.32}$$

where $\pi_d^{(t-1)} = \Pi_{t-1}(\boldsymbol{\Theta}^{(c,d)})$, $\forall c \in \mathcal{C}$, $d \in \mathcal{D}_c$. Therein, the summation $\sum_{c \in \Delta}$ is taken over $\Delta \in \mathscr{P}(\mathcal{C})$ that satisfies $c \in \Delta$, and $\mathscr{P}(\mathcal{C})$ is the power set of $\mathcal{C}$ consisting of all subsets of $\mathcal{C}$.

Lemma EC.3 is technical, which completely characterizes the sampling policy.

Proof. Equality (EC.32) follows from the definition of ψ_t naturally. Let $\hat{\boldsymbol{\theta}}_{2,k}^{(t)}$, $\forall k \geq 1$ denote the independent copies of $\hat{\boldsymbol{\theta}}_2^{(t)}$ generated in Algorithm 1, and $\hat{d}_{2,k}^{(t)}(c) = \arg\max_{d \in \mathcal{D}_c} (\hat{\mu}_{2,k}^{(t)})_d$, accordingly. We define an array of stopping times by $\kappa_c = \min\{k \geq 1 : \hat{d}_{2,k}^{(t)}(c) \neq \hat{d}_1^{(t)}(c)\}$. Conditioned on $\mathcal{E}_{t-1}$ and $\hat{\boldsymbol{d}}_1$, each κ_c follows an independent geometric distribution with parameter $\left(1-\pi_{\hat{d}_1^{(t)}(c)}^{(t-1)}\right)$. By construction, for all $\Delta \in \mathscr{P}(\mathcal{C})$,

$$\begin{aligned}\mathbb{P}\left(\Delta_t = \Delta \middle| \mathcal{E}_{t-1}, \hat{\boldsymbol{d}}_1\right) &= \sum_{\tau=1}^{\infty} \mathbb{P}\left(\kappa_{c'} = \tau, \forall c' \in \Delta,\, \kappa_{c''} > \tau, \forall c'' \notin \Delta \middle| \mathcal{E}_{t-1}, \hat{\boldsymbol{d}}_1\right) \\ &= \sum_{\tau=1}^{\infty} \prod_{c' \in \Delta} \left(\pi_{\hat{d}_1^{(t)}(c')}^{(t-1)}\right)^{\tau-1} \left(1-\pi_{\hat{d}_1^{(t)}(c')}^{(t-1)}\right) \prod_{c' \notin \Delta} \left(\pi_{\hat{d}_1^{(t)}(c')}^{(\tau-1)}\right)^{\tau} \\ &= \frac{\prod_{c \in \Delta}\left(1-\pi_{\hat{d}_1^{(t)}(c)}^{(t-1)}\right)\prod_{c \notin \Delta} \pi_{\hat{d}_1^{(t)}(c)}^{(t-1)}}{1-\prod_{c \in \mathcal{C}} \pi_{\hat{d}_1^{(t)}(c)}^{(t-1)}},\end{aligned}$$

where the first equality follows from the definition of Δ_t, which consists of contexts that first yield estimates indicating a best design different from the initially estimated one during the simulation process. The second equality follows from the fact that κ_c are geometrically distributed, and the last equality follows from straightforward calculations.

Therefore, using the law of total probability by conditioning on Δ_t, we have

$$\begin{aligned}\mathbb{P}\left(C_t = c \middle| \mathcal{E}_{t-1}, \hat{\boldsymbol{d}}_1\right) &= \sum_{\Delta \in \mathscr{P}(\mathcal{C})} \mathbb{P}\left(\Delta_t = \Delta \middle| \mathcal{E}_{t-1}, \hat{\boldsymbol{d}}_1\right) \cdot \mathbb{P}\left(C_t = c \middle| \mathcal{E}_{t-1}, \hat{\boldsymbol{d}}_1, \Delta_t = \Delta\right) \\ &= \sum_{\Delta \in \mathscr{P}(\mathcal{C})} \mathbb{P}\left(\Delta_t = \Delta \middle| \mathcal{E}_{t-1}, \hat{\boldsymbol{d}}_1\right) \cdot \frac{\mathbf{1}\{c \in \Delta\}}{|\Delta|} \\ &= \sum_{\Delta \in \mathscr{P}(\mathcal{C}): c \in \Delta} \frac{\prod_{c' \in \Delta}\left(1-\pi_{\hat{d}_1^{(t)}(c')}^{(t-1)}\right)\prod_{c' \notin \Delta} \pi_{\hat{d}_1^{(t)}(c')}^{(t-1)}}{1-\prod_{c' \in \mathcal{C}} \pi_{\hat{d}_1^{(t)}(c')}^{(t-1)}} \frac{1}{|\Delta|}.\end{aligned}$$

Therefore, (EC.31) follows immediately from

$$\alpha_t(c) = \mathbb{P}\left(C_t = c | \mathcal{E}_{t-1}\right) = \mathbb{E}\left[\mathbb{P}\left(C_t = c \middle| \mathcal{E}_{t-1}, \hat{\boldsymbol{d}}_1\right) \middle| \mathcal{E}_{t-1}\right].$$

Then it suffices to show (EC.30). Note that for each $c \in \mathcal{C}$, the estimated second best design $\hat{d}_{2,k}^{(t)}(c)$ in context c follows a categorical distribution over $\mathcal{D}_c$, and that for $d \in \mathcal{D}_c$ with $d \neq \hat{d}_1^{(t)}(c)$,

$$\mathbb{P}\left(\hat{d}_{2,\kappa_c}^{(t)}(c) = d \middle| \mathcal{E}_{t-1}, \hat{\boldsymbol{d}}_1, \kappa_c\right) = \pi_d^{(t-1)} \Big/ \left(1 - \pi_{\hat{d}_1^{(t)}(c)}^{(t-1)}\right).$$

This implies that conditioned on $\mathcal{E}_{t-1}$ and $\hat{\boldsymbol{d}}_1$, the random variable $\hat{d}_{2,\kappa_c}^{(t)}(c)$ is independent of κ_c, and thus independent of C_t. Therefore, for each $c \in \mathcal{C}$ and $d \in \mathcal{D}_c$ with $d \neq \hat{d}_1^{(t)}(c)$, we have

$$\mathbb{P}\left(\hat{d}_{2,\kappa_c}^{(t)}(c) = d \middle| \mathcal{E}_{t-1}, \hat{\boldsymbol{d}}_1\right) = \frac{\pi_d^{(t-1)}}{\left(1 - \pi_{\hat{d}_1^{(t)}(c)}^{(t-1)}\right)},$$

and

$$\begin{aligned}
\mathbb{P}\left(D_t = d \middle| \mathcal{E}_{t-1}, \hat{\boldsymbol{d}}_1\right) =& \mathbb{P}\left(C_t = c, D_t = d \middle| \mathcal{E}_{t-1}, \hat{\boldsymbol{d}}_1\right) \\
=& \mathbb{P}\left(C_t = c \middle| \mathcal{E}_{t-1}, \hat{\boldsymbol{d}}_1\right) \cdot \mathbf{1}\{\hat{d}_1^{(t)}(c) = d\} \cdot \gamma(c) \\
&+ \mathbb{P}\left(C_t = c, \hat{d}_{2,\kappa_c}^{(t)}(c) = d \middle| \mathcal{E}_{t-1}, \hat{\boldsymbol{d}}_1\right) \cdot \mathbf{1}\{\hat{d}_1^{(t)}(c) \neq d\} \cdot (1 - \gamma(c)) \\
=& \mathbb{P}\left(C_t = c \middle| \mathcal{E}_{t-1}, \hat{\boldsymbol{d}}_1\right) \cdot \Big[\mathbf{1}\{\hat{d}_1^{(t)}(c) = d\} \cdot \gamma(c) \\
&+ \mathbb{P}\left(\hat{d}_{2,\kappa_c}^{(t)}(c) = d \middle| \mathcal{E}_{t-1}, \hat{\boldsymbol{d}}_1\right) \cdot \mathbf{1}\{\hat{d}_1^{(t)}(c) \neq d\} \cdot (1 - \gamma(c))\Big],
\end{aligned}$$

where the last equality follows from the conditional independence of $\hat{d}_{2,\kappa_c}^{(t)}(c)$ and C_t. Then (EC.30) follows immediately from the two equations above and

$$\psi_t(c, d) = \mathbb{P}\left(D_t = d | \mathcal{E}_{t-1}\right) = \mathbb{E}\left[\mathbb{P}\left(D_t = d \middle| \mathcal{E}_{t-1}, \hat{\boldsymbol{d}}_1\right) \middle| \mathcal{E}_{t-1}\right]. \quad \square$$

To shed some light on the (conditional) sampling ratios of TTTS-C in each context, we focus on a fixed context c and a design $d \in \mathcal{D}_c$. We can decompose the summations in Equation (EC.30) and (EC.31) over functions $d_1 : \mathcal{C} \to \mathcal{D}$ as $\sum_{\boldsymbol{d}_1} \cdot = \sum_{d_1(-c)} \sum_{d_1(c) \in \mathcal{D}_c} \cdot$, where $d_1(-c)$ denotes that the first summation is taken over all combinations of the function d_1 evaluated at $\mathcal{C} \backslash \{c\}$. Given the value of $d_1(-c)$, the corresponding summands in (EC.30) and (EC.31) can be rewritten as follows by extracting common factors (in the summation $\sum_{d_1(c) \in \mathcal{D}_c} \cdot$) that are irrelevant to the value of $d_1(c)$:

$$J_{d_1(-c)}^{(t-1)} = A_{d_1(-c)}^{(t-1)} \left(\gamma(c) \frac{\pi_d^{(t-1)}\left(1 - \pi_d^{(t-1)}\right)}{1 - B_{d_1(-c)}^{(t-1)} \cdot \pi_d^{(t-1)}} + (1 - \gamma(c)) \sum_{d_1(c) \neq d} \frac{\pi_d^{(t-1)} \pi_{d_1(c)}^{(t-1)}}{1 - B_{d_1(-c)}^{(t-1)} \cdot \pi_{d_1(c)}^{(t-1)}}\right) \tag{EC.33}$$

and

$$K_{d_1(-c)}^{(t-1)} = A_{d_1(-c)}^{(t-1)} \sum_{d_1(c) \in \mathcal{D}_c} \frac{\pi_{d_1(c)}^{(t-1)}\left(1 - \pi_{d_1(c)}^{(t-1)}\right)}{1 - B_{d_1(-c)}^{(t-1)} \cdot \pi_{d_1(c)}^{(t-1)}}, \tag{EC.34}$$

respectively. Therein, $A^{(t-1)}_{d_1(-c)}$ is the extracted common factor given by

$$A^{(t-1)}_{d_1(-c)} = \sum_{c\in\Delta} B^{(t-1)}_{d_1(-c)} \cdot \prod_{c'\in\Delta\setminus\{c\}} \left(1-\pi^{(t-1)}_{d_1(c')}\right) \prod_{c'\notin\Delta} \pi^{(t-1)}_{d_1(c')} \frac{1}{|\Delta|},$$

and moreover,

$$B^{(t-1)}_{d_1(-c)} = \prod_{c'\in\mathcal{C}\setminus\{c\}} \pi^{(t-1)}_{d_1(c')}.$$

Then $\psi_t(d) = \sum_{d_1(-c)} J^{(t-1)}_{d_1(-c)}$ and $\alpha_t(c) = \sum_{d_1(-c)} K^{(t-1)}_{d_1(-c)}$.

Notice that $A^{(t-1)}_{d_1(-c)}$ and $B^{(t-1)}_{d_1(-c)}$ are irrelevant to the posterior distribution of parameters associated with context c. For $d\neq d^*(c)$, simple calculation yields that the ratio of (EC.33) over (EC.34) is asymptotically equivalent to $\pi^{(t-1)}_d/(1-\pi^{(t-1)}_{d^*(c)})$. Thus, a sub-optimal design in each context is allocated a minimum conditional sampling effort proportional to its posterior probability of being the best design. On the other hand, the ratio of (EC.33) over (EC.34) can be lower and upper bounded by $\gamma(c)/(1+\delta_t)$ and $\gamma(c)\cdot(1+\delta_t)$, respectively, where $\delta_t = \mathcal{O}(|1/\pi^{(t-1)}_d - B^{(t-1)}_{d_1(-c)}|)$. In particular, for $d=d^*(c)$, if $B^{(t-1)}_{d_1(-c)}\to 1$ and $\pi^{(t-1)}_{d^*(c)}\to 1$, then the ratio will approach $\gamma(c)$, hence so will $\beta_t(c,d^*(c))$. This suggests that the conditional sampling effort allocated to contextual best designs is approximately determined by the hyperparameter $\boldsymbol{\gamma}$. The above discussions are only a sketch of the main ideas and do not constitute a rigorous proof, as $\beta_t(c,d^*(c))$ is the ratio of summations of (EC.33) over (EC.34), and the weight $B^{(t-1)}_{d_1(-c)}$ does not tend to 1 for each term.

Next, we present the rigorous proof of Lemma 1.

Proof of Lemma 1. We first prove the second statement. Fix $c\in\mathcal{C}$ and $d\neq d^*(c)$. For any fixed value of $d_1(-c)$, we have a lower bound of (EC.33)

$$J^{(t-1)}_{d_1(-c)} \geq A^{(t-1)}_{d_1(-c)}(1-\gamma(c))\frac{\pi^{(t-1)}_d\pi^{(t-1)}_{d^*(c)}}{1-B^{(t-1)}_{d_1(-c)}\cdot\pi^{(t-1)}_{d^*(c)}}. \tag{EC.35}$$

On the other hand, (EC.34) can be upper bounded by

$$K^{(t-1)}_{d_1(-c)} \leq A^{(t-1)}_{d_1(-c)}|\mathcal{D}_c| \max_{d'\in\mathcal{D}_c\setminus\{d^*(c)\}} \frac{\pi^{(t-1)}_{d'}\left(1-\pi^{(t-1)}_{d'}\right)}{1-B^{(t-1)}_{d_1(-c)}\cdot\pi^{(t-1)}_{d'}} = A^{(t-1)}_{d_1(-c)}|\mathcal{D}_c|\frac{\pi^{(t-1)}_{d^*(c)}\left(1-\pi^{(t-1)}_{d^*(c)}\right)}{1-B^{(t-1)}_{d_1(-c)}\cdot\pi^{(t-1)}_{d^*(c)}}. \tag{EC.36}$$

In fact, notice that the numerical inequality $x(1-x)\geq y(1-y)$ for $1/2<x<1$ and $0<y\leq 1-x$, and the observation $\sum_{d'\in\mathcal{D}_c\setminus\{d^*(c)\}}\pi^{(t-1)}_{d'} = 1-\pi^{(t-1)}_{d^*(c)}$ implies that $\pi^{(t-1)}_{d'}(1-\pi^{(t-1)}_{d'}) \leq \pi^{(t-1)}_{d^*(c)}(1-\pi^{(t-1)}_{d^*(c)})$ for $d'\in\mathcal{D}_c\setminus\{d^*(c)\}$ holds true, because $\pi^{(t-1)}_{d^*(c)}\geq\frac{1}{2}$ by assumption. Also, notice that $\pi^{(t-1)}_{d'}\leq\frac{1}{2}\leq\pi^{(t-1)}_{d^*(c)}$ holds true, hence $1-B^{(t-1)}_{d_1(-c)}\pi^{(t-1)}_{d'}\geq 1-B^{(t-1)}_{d_1(-c)}\pi^{(t-1)}_{d^*(c)}$. Therefore, the maximum in (EC.36) is attained at $d'=d^*(c)$.

Combining (EC.35) and (EC.36),

$$\frac{J^{(t-1)}_{d_1(-c)}}{K^{(t-1)}_{d_1(-c)}} \geq \frac{1-\gamma(c)}{|\mathcal{D}_c|}\frac{\pi^{(t-1)}_d}{1-\pi^{(t-1)}_{d^*(c)}},$$

which is irrelevant to $d_1(-c)$. Hence following (EC.32),

$$\beta_t(c,d) = \frac{\sum_{d_1(-c)} J^{(t-1)}_{d_1(-c)}}{\sum_{d_1(-c)} K^{(t-1)}_{d_1(-c)}} \geq \frac{1-\gamma(c)}{|\mathcal{D}_c|}\frac{\pi^{(t-1)}_d}{1-\pi^{(t-1)}_{d^*(c)}}.$$

Similarly, we have

$$J_{d_1(-c)}^{(t-1)} \leq A_{d_1(-c)}^{(t-1)} \frac{\pi_d^{(t-1)}}{1 - B_{d_1(-c)}^{(t-1)} \cdot \pi_{d^*(c)}^{(t-1)}},$$

and

$$K_{d_1(-c)}^{(t-1)} \geq A_{d_1(-c)}^{(t-1)} \frac{\pi_{d^*(c)}^{(t-1)} \left(1 - \pi_{d^*(c)}^{(t-1)}\right)}{1 - B_{d_1(-c)}^{(t-1)} \cdot \pi_{d^*(c)}^{(t-1)}}.$$

Therefore,

$$\beta_t(c,d) = \frac{\sum_{d_1(-c)} J_{d_1(-c)}^{(t-1)}}{\sum_{d_1(-c)} K_{d_1(-c)}^{(t-1)}} \leq \frac{\pi_d^{(t-1)}}{\pi_{d^*(c)}^{(t-1)} \left(1 - \pi_{d^*(c)}^{(t-1)}\right)} \leq 2 \frac{\pi_d^{(t-1)}}{1 - \pi_{d^*(c)}^{(t-1)}},$$

which completes the proof for the second statement. Now we prove the first statement.

For $c \in \mathcal{C}$ and $d = d^*(c) \in \mathcal{D}_c$, we first focus on a particular value of $d_1(-c)$, which coincides with $d^*(-c)$, i.e., $d_1(c') = d^*(c')$ for all $c' \in \mathcal{C} \backslash \{c\}$. We have

$$\frac{\pi_{d'}^{(t-1)} \cdot \pi_{d^*(c)}^{(t-1)}}{1 - B_{d_1(-c)}^{(t-1)} \cdot \pi_{d'}^{(t-1)}} \leq \frac{\pi_{d'}^{(t-1)} \cdot \pi_{d^*(c)}^{(t-1)}}{1 - \pi_{d'}^{(t-1)}} \leq \frac{(1 - \pi_{d^*(c)}^{(t-1)}) \cdot \pi_{d^*(c)}^{(t-1)}}{1 - (1 - \pi_{d^*(c)}^{(t-1)})} = 1 - \pi_{d^*(c)}^{(t-1)}, \tag{EC.37}$$

where the second inequality follows from $\pi_{d'}^{(t-1)} \leq \sum_{d' \in \mathcal{D}_c} \pi_{d'}^{(t-1)} = 1 - \pi_{d'}^{(t-1)}$ and thus $1 - \pi_{d'}^{(t-1)} \geq 1 - (1 - \pi_{d'}^{(t-1)})$. Now, (EC.33) can be lower bounded by

$$J_{d_1(-c)}^{(t-1)} \geq A_{d_1(-c)}^{(t-1)} \gamma(c) \frac{\pi_{d^*(c)}^{(t-1)} \left(1 - \pi_{d^*(c)}^{(t-1)}\right)}{1 - B_{d_1(-c)}^{(t-1)} \cdot \pi_{d^*(c)}^{(t-1)}}, \tag{EC.38}$$

and upper bounded by

$$\begin{aligned} J_{d_1(-c)}^{(t-1)} &\leq A_{d_1(-c)}^{(t-1)} \left(\gamma(c) \frac{\pi_{d^*(c)}^{(t-1)} \left(1 - \pi_{d^*(c)}^{(t-1)}\right)}{1 - B_{d_1(-c)}^{(t-1)} \cdot \pi_{d^*(c)}^{(t-1)}} + (1 - \gamma(c)) |\mathcal{D}_c| \max_{d' \in \mathcal{D}_c \backslash \{d^*(c)\}} \frac{\pi_{d'}^{(t-1)} \cdot \pi_{d^*(c)}^{(t-1)}}{1 - B_{d_1(-c)}^{(t-1)} \cdot \pi_{d'}^{(t-1)}} \right) \\ &\leq A_{d_1(-c)}^{(t-1)} \left(\gamma(c) \frac{\pi_{d^*(c)}^{(t-1)} \left(1 - \pi_{d^*(c)}^{(t-1)}\right)}{1 - B_{d_1(-c)}^{(t-1)} \cdot \pi_{d^*(c)}^{(t-1)}} + (1 - \gamma(c)) |\mathcal{D}_c| \left(1 - \pi_{d^*(c)}^{(t-1)}\right) \right), \end{aligned} \tag{EC.39}$$

where the last inequality follows from (EC.37). Similarly, we can bound (EC.34), for sufficiently large t, by

$$K_{d_1(-c)}^{(t-1)} \geq A_{d_1(-c)}^{(t-1)} \frac{\pi_{(c,d^*(c))}^{(t-1)} \left(1 - \pi_{(c,d^*(c))}^{(t-1)}\right)}{1 - B_{d_1(-c)}^{(t-1)} \cdot \pi_{(c,d^*(c))}^{(t-1)}}, \tag{EC.40}$$

and

$$K_{d_1(-c)}^{(t-1)} \leq A_{d_1(-c)}^{(t-1)} \left(\frac{\pi_{d^*(c)}^{(t-1)} \left(1 - \pi_{d^*(c)}^{(t-1)}\right)}{1 - B_{d_1(-c)}^{(t-1)} \cdot \pi_{d^*(c)}^{(t-1)}} + \left(1 - \pi_{d^*(c)}^{(t-1)}\right) \right). \tag{EC.41}$$

Combining (EC.38)-(EC.41), we see that

$$\gamma(c) / (1 + \delta_t) \leq \frac{J_{d_1(-c)}^{(t-1)}}{K_{d_1(-c)}^{(t-1)}} \leq \gamma(c) \cdot (1 + \delta_t), \tag{EC.42}$$

where

$$\delta_t = |\mathcal{D}_c| \frac{1 - B^{(t-1)}_{d_1(-c)} \cdot \pi^{(t-1)}_{d^*(c)}}{\pi^{(t-1)}_{d^*(c)}}.$$

To see δ_t is a vanishing value, notice that we have now focused on the particular value $d_1(-c) = d^*(-c)$. Also recall that $B^{(t-1)}_{d_1(-c)} = B^{(t-1)}_{d^*(-c)} = \prod_{c' \in \mathcal{C} \setminus \{c\}} \pi^{(t-1)}_{d^*(c')}$. By assumption $\pi^{(t-1)}_{d^*(c)} \to 1$, for any $0 < \varepsilon < 1$, there exists an N (not necessarily deterministic) such that for each $t \geq N$ and $c' \in \mathcal{C}$, we have

$$\pi^{(t-1)}_{d^*(c)} \geq 1 - \varepsilon/2|\mathcal{C}|^{|\mathcal{D}|}. \tag{EC.43}$$

Then by the numerical inequality $1/(1-x) - (1-x) \leq 3x$ for $0 \leq x \leq 1/2$, we see that $\delta_t \leq 3\varepsilon$.

Moreover, $A^{(t-1)}_{d_1(-c)}, B^{(t-1)}_{d_1(-c)} \leq \varepsilon/2|\mathcal{C}|^{|\mathcal{D}|}$ for any $d_1(-c) \neq d^*(-c)$. It follows from (EC.43) that for all $d_1(-c) \neq d^*(-c)$, we have $B^{(t-1)}_{d_1(-c)} < B^{(t-1)}_{d^*(-c)}$. By definitions (EC.33) and (EC.34), $J^{(t-1)}_{d_1(-c)}/A^{(t-1)}_{d_1(-c)}$ and $K^{(t-1)}_{d_1(-c)}/A^{(t-1)}_{d_1(-c)}$ are increasing functions of $B^{(t-1)}_{d_1(-c)}$. Therefore, $J^{(t-1)}_{d_1(-c)}/A^{(t-1)}_{d_1(-c)} \leq J^{(t-1)}_{d^*(-c)}/A^{(t-1)}_{d^*(-c)}$, and $K^{(t-1)}_{d_1(-c)}/A^{(t-1)}_{d_1(-c)} \leq K^{(t-1)}_{d^*(-c)}/A^{(t-1)}_{d^*(-c)}$. By rearranging the terms, we see that

$$\frac{J^{(t-1)}_{d_1(-c)}}{J^{(t-1)}_{d^*(-c)}} \leq \frac{A^{(t-1)}_{d_1(-c)}}{A^{(t-1)}_{d^*(-c)}} \leq \frac{\varepsilon}{|\mathcal{C}|^{|\mathcal{D}|-1}}, \quad \text{and} \quad \frac{K^{(t-1)}_{d_1(-c)}}{K^{(t-1)}_{d^*(-c)}} \leq \frac{A^{(t-1)}_{d_1(-c)}}{A^{(t-1)}_{d^*(-c)}} \leq \frac{\varepsilon}{|\mathcal{C}|^{|\mathcal{D}|-1}}. \tag{EC.44}$$

Combining (EC.42)-(EC.44), we see that

$$\beta_t(c, d^*(c)) = \frac{J^{(t-1)}_{d^*(-c)} + \sum_{d_1(-c) \neq d^*(-c)} J^{(t-1)}_{d_1(-c)}}{K^{(t-1)}_{d^*(-c)} + \sum_{d_1(-c) \neq d^*(-c)} K^{(t-1)}_{d_1(-c)}} \leq \gamma(c) \cdot (1 + \delta_t)(1 + \varepsilon) \leq \gamma(c)(1 + 5\varepsilon).$$

This implies that $\limsup_{t\to\infty} \beta_t(c, d^*(c)) \leq \gamma(c)$. Similarly, we see that

$$\beta_t(c, d^*(c)) \geq \gamma(c)/(1 + 5\varepsilon),$$

and thus $\liminf_{t\to\infty} \beta_t(c, d^*(c)) \geq \gamma(c)$, which completes the proof. □

The next proof establishes Lemma 2 for the aggregate sampling ratio for each context.

Proof of Lemma 2. Fix a context $c \in \mathcal{C}$. By Lemma EC.3, the sampling effort for context c can be lower bounded by relaxing the summation in (EC.31) to the dominating term with $\boldsymbol{d}_1 = \boldsymbol{d}^*$ and $\Delta = \{c\}$, i.e., with probability 1, for sufficiently large t,

$$\alpha_t(c) \geq \frac{\prod_{c' \in \mathcal{C}} \pi^{(t-1)}_{d^*(c')}}{1 - \prod_{c' \in \mathcal{C}} \pi^{(t-1)}_{d^*(c')}} \left(1 - \pi^{(t-1)}_{d^*(c)}\right) \prod_{c' \in \mathcal{C} \setminus \{c\}} \pi^{(t-1)}_{d^*(c')} \geq \frac{1}{2} \frac{1 - \pi^{(t-1)}_{d^*(c)}}{1 - \prod_{c' \in \mathcal{C}} \pi^{(t-1)}_{d^*(c')}}, \tag{EC.45}$$

where the last inequality follows from Theorem 3. The second inequality in (3) follows immediately from the second term in (EC.44) by choosing ε to be small enough.

It remains to show (4). By taking $\tilde{\boldsymbol{\Theta}} = \bigcup_{d' \in \mathcal{D}_c} \{\mu_{d'} > \mu_{d^*(c)}\}$ in Theorem 1, combining (EC.45), we have

$$\begin{aligned} \alpha_t(c) &\geq \frac{1}{2} \exp\left\{ -(t-1) \left(\bar{\alpha}_{t-1}(c) \min_{d \in \mathcal{D}_c \setminus \{d^*(c)\}} G_d(\bar{\beta}_{t-1}(c, d^*(c)), \bar{\beta}_{t-1}(c, d)) + \varepsilon_{t-1} \right) \right\} \\ &\geq \frac{1}{2} \exp\left\{ -2\Gamma^* \cdot (t-1)\bar{\alpha}_{t-1}(c) \right\}, \end{aligned}$$

where ε_t is a vanishing sequence as $t \to 0$. Note the inequality that

$$\begin{aligned} T\bar{\alpha}_T(c) = \sum_{t=1}^{T} \alpha_t(c) &= (T-1)\bar{\alpha}_{T-1}(c) + \alpha_T(c) \\ &\geq (T-1)\bar{\alpha}_{T-1}(c) + \frac{1}{2} \exp\left\{ -2\Gamma^* \cdot (T-1)\bar{\alpha}_{T-1}(c) \right\}. \end{aligned} \tag{EC.46}$$

Note that $T\bar{\alpha}_T(c)$ is non-decreasing with respect to T. Therefore, we have either $T\bar{\alpha}_T(c)\to\infty$ or $\lim_{T\to\infty}T\bar{\alpha}_T(c)=:L<\infty$. The latter immediately leads to a contradiction by $\frac{1}{2}\exp\{-2L\Gamma^*\}\le T\bar{\alpha}_T(c)-(T-1)\bar{\alpha}_{T-1}(c)\to 0$. Now, the derivative of the right-hand side of (EC.46) with respect to $(T-1)\bar{\alpha}_{T-1}(c)$, i.e., $1-\Gamma^*\exp\{-2\Gamma^*\cdot(T-1)\bar{\alpha}_{T-1}(c)\}$, is non-negative for T large enough since $(T-1)\bar{\alpha}_{T-1}(c)\to\infty$.

Let $\tilde{C}$ be a random real to be determined. If $T\bar{\alpha}_T(c)\ge\frac{1}{2\Gamma^*}\ln\left(\tilde{C}+\Gamma^*\cdot(T-1)\right)$, then by (EC.46),

$$(T+1)\bar{\alpha}_{T+1}(c)\ge\frac{1}{2\Gamma^*}\ln\left(\tilde{C}+\Gamma^*\cdot(T-1)\right)+\frac{1}{2}\cdot\frac{1}{\tilde{C}+\Gamma^*\cdot(T-1)}\ge\frac{1}{2\Gamma^*}\ln\left(\tilde{C}+\Gamma^*\cdot T\right),$$

where the last inequality follows from the numerical inequality $x\ge\ln(1+x)$ with $x=\Gamma^*/\left(\tilde{C}+\Gamma^*\cdot(T-1)\right)$. Now it suffices to construct the random variable $\tilde{C}$ and apply the mathematical induction. Let $\tilde{T}$ be the smallest number such that all of the aforementioned asymptotic statements hold. Then we complete the proof by setting $\tilde{C}=\exp\{2\Gamma^*\cdot\tilde{T}\bar{\alpha}_{\tilde{T}}(c)\}-\Gamma^*\cdot(\tilde{T}-1)$. ☐

D.2. Proof for $m_c=1$. We first show that TTTS-C for selecting contextual best designs asymptotically satisfies the balance condition (12).

THEOREM EC.2 **(Theorem 2 (restated))**. *Suppose Assumptions 1 and 2 (or 2′) hold. For TTTS-C with hyperparameter $\boldsymbol{\gamma}>0$ in Algorithm 1, the balance condition (10) is asymptotically satisfied, i.e., there exists $z\in\mathbb{R}^+$, such that for all $c\in\mathcal{C}$, $d\in\mathcal{D}_c\backslash\{d^*(c)\}$, we have*

$$\lim_{T\to\infty}\bar{\alpha}_T(c)G_d(\gamma(c),\bar{\beta}_T(c,d))=z.$$

Immediately we have the following corollary.

COROLLARY EC.2 **(Corollary 2 (restated; Asymptotic Quasi-Optimality))**. *Suppose Assumptions 1 and 2 (or 2′) hold. For TTTS-C with hyperparameter $\boldsymbol{\gamma}>0$ in Algorithm 1,*

$$\lim_{T\to\infty}-\frac{1}{T}\ln\Pi_T\left(\bigcup_{c\in\mathcal{C}}\bigcup_{d\in\mathcal{D}\backslash\{d^*(c)\}}\boldsymbol{\Theta}_{(c,d)}\right)=\Gamma^*_{\boldsymbol{\gamma}}.$$

For any sequential sampling policy with $\lim_{T\to\infty}\bar{\beta}_T(c,d^*(c))=\gamma(c)$,

$$\limsup_{T\to\infty}-\frac{1}{T}\ln\Pi_T\left(\bigcup_{c\in\mathcal{C}}\bigcup_{d\in\mathcal{D}\backslash\{d^*(c)\}}\boldsymbol{\Theta}_{(c,d)}\right)\le\Gamma^*_{\boldsymbol{\gamma}}.$$

Proof. Consider any subsequences $\{\bar{\alpha}_{T_n}\}_{n\ge1}$ and $\{\bar{\beta}_{T_n}\}_{n\ge1}$ of $\{\bar{\alpha}_T\}_{T\ge1}$ and $\{\bar{\beta}_T\}_{T\ge1}$, respectively. By Theorem 2, there exists $z\in\mathbb{R}^+$ such that

$$\lim_{n\to\infty}\bar{\alpha}_{T_n}(c)G_d(\bar{\beta}_{T_n}(c,d^*(c)),\bar{\beta}_{T_n}(c,d))-z=0,\quad\forall c\in\mathcal{C},d\in\mathcal{D}_c. \tag{EC.47}$$

Since $\bar{\alpha}_T$ and $\bar{\beta}_T$ are bounded, there exist cluster points α_0 and β_0, and further subsequences $\{\bar{\alpha}_{T_{n_k}}\}_{k\ge1}$ and $\{\bar{\beta}_{T_{n_k}}\}_{k\ge1}$ that converge to the cluster points, respectively. By Lemma 1, we have $\beta_0(c,d^*(c))=\gamma(c)$, $\forall c\in\mathcal{C}$. Therefore, (α_0,β_0) is feasible for the optimization problem (11). Since $G_d(\cdot,\cdot)$ is continuous for all $d\in\mathcal{D}$, it follows from (EC.47) that $\alpha_0(c)G_d(\gamma(c),\beta_0(c,d))=z$, $\forall c\in\mathcal{C},d\in\mathcal{D}_c\backslash\{d^*(c)\}$. According to Proposition 3, (α_0,β_0) is optimal to (11). In other words, for $c\in\mathcal{C}$, $d\in\mathcal{D}_c\backslash\{d^*(c)\}$,

$$\lim_{k\to\infty}\bar{\alpha}_{T_{n_k}}(c)G_d(\bar{\beta}_{T_{n_k}}(c,d^*(c)),\bar{\beta}_{T_{n_k}}(c,d))=\alpha_0(c)G_d(\gamma(c),\beta_0(c,d))=\Gamma^*_{\boldsymbol{\gamma}}.$$

As a consequence of Theorem 1,

$$\lim_{k\to\infty} -\frac{1}{T_{n_k}} \ln \Pi_{T_{n_k}} \left(\bigcup_{c\in\mathcal{C}} \bigcup_{d\in\mathcal{D}\setminus\{d^*(c)\}} \mathbf{\Theta}_{(c,d)} \right) = \lim_{k\to\infty} \min_{c\in\mathcal{C}} \min_{d\in\mathcal{D}_c\setminus\{d^*\}} \bar{\alpha}_{T_{n_k}}(c) G_d(\bar{\beta}_{T_{n_k}}(c,d^*(c)), \bar{\beta}_{T_{n_k}}(c,d)) \\ = \Gamma^*_{\boldsymbol{\gamma}}.$$

That is, each subsequence has a further convergent subsequence that converges to $\Gamma^*_{\boldsymbol{\gamma}}$, which leads to the first statement. The second statement is a direct consequence of Corollary 1 and the definition of $\Gamma^*_{\boldsymbol{\gamma}}$. □

Then we return to show the main result.

Proof of Theorem 2. This proof consists of three steps.

Step 1: Consistency. By Theorem 3 and Lemma EC.1, for each context c and each design $d \in \mathcal{D}_c\setminus\{d^*(c)\}$, and for $\varepsilon = \min_{c\in\mathcal{C}} \min_{d',d''} \in \mathcal{D}_c : d' \neq d'' |\mu_{d'} - \mu_{d''}|/2$, as $t\to\infty$,

$$\pi_d^{(t)} = \Pi_t(\mathbf{\Theta}_{(c,d)}) \le \Pi_t(\{\boldsymbol{\theta}\in\mathbf{\Theta} \big| |\mu_d - \mu_d^*| \ge \varepsilon\}) + \Pi_t(\{\boldsymbol{\theta}\in\mathbf{\Theta} \big| |\mu_{d^*(c)} - \mu^*_{d^*(c)}| \ge \varepsilon\}) \to 0.$$

Therefore, the posterior probability of a sub-optimal design being the context-dependent best is diminishing. As a consequence, Lemma 1 is valid.

Step 2: Local balance of sampling ratios in a fixed context c.

Step 2.1. Define $z_c := \Gamma^*_{\gamma(c)}$. We claim that, with probability 1,

$$\limsup_{T\to\infty} G_d(\gamma(c), \bar{\beta}_T(c,d)) - z_c \le 0, \quad \forall d\in\mathcal{D}_c\setminus\{d^*(c)\}. \tag{EC.48}$$

We first show that $\liminf_{T\to\infty} \left(G_d(\gamma(c), \bar{\beta}_T(c,d)) - z_c\right)$ is non-positive. Suppose otherwise that the limit inferior is $\delta > 0$. Then by Lemma 1, with probability 1, for sufficiently large T,

$$\beta_T(c,d) \le 2\frac{\pi_d^{(T-1)}}{1-\pi^{(T-1)}_{d^*(c)}} \le 2\frac{\pi_d^{(T-1)}}{\max_{d'\in\mathcal{D}_c\setminus\{d^*(c)\}} \pi_{d'}^{(T-1)}}. \tag{EC.49}$$

According to the second statement of Theorem 1, there exists a vanishing sequence of ε_T such that $\varepsilon_T \to 0$, and with probability 1, for all $d'\in\mathcal{D}_c$ and sufficiently large T,

$$\begin{aligned} &\exp\{-(T-1)\bar{\alpha}_{T-1}(c)\left(G_{d'}(\gamma(c), \bar{\beta}_{T-1}(c,d')) + \varepsilon_{T-1}\right)\} \le \pi_{d'}^{(T-1)} \\ \le &\exp\{-(T-1)\bar{\alpha}_{T-1}(c)\left(G_{d'}(\gamma(c), \bar{\beta}_{T-1}(c,d')) - \varepsilon_{T-1}\right)\}. \end{aligned} \tag{EC.50}$$

With a slight abuse of notation, we will refer to $\{\varepsilon_T\}_{T\ge1}$ as several different vanishing sequences introduced below. Since the maximum of a finite number of vanishing sequences also vanishes, this notion can be formalized by redefining $\{\varepsilon_T\}_{T\ge1}$ as maximum of these sequences, while we omit these tedious statements.

Now, let $\Gamma^*_{\beta_T(c)}$ be the optimal value defined in optimization problem (13) with $\gamma(c)$ replaced by $\beta_T(c) := \bar{\beta}_T(c, d^*(c))$. Then by definition, we have

$$\min_{d'\in\mathcal{D}_c\setminus\{d^*(c)\}} G_d(\beta_T(c), \bar{\beta}_T(c,d')) \le \Gamma^*_{\beta_T(c)}.$$

According to Lemma 1, we see that $\beta_T(c)\to\gamma(c)$, and $\Gamma^*_{\beta_T(c)} \to \Gamma^*_{\gamma(c)}$. Note that $G_d(\cdot,\cdot)$ is a concave function on a compact set, and thus is uniformly continuous. As a result, there exists a vanishing sequence $\{\tilde{\varepsilon}_T\}_{T\ge1}$ such that

$$\min_{d'\in\mathcal{D}_c\setminus\{d^*(c)\}} G_d(\gamma(c), \bar{\beta}_T(c,d')) \le \Gamma^*_{\gamma(c)} + \tilde{\varepsilon}_T = z_c + \tilde{\varepsilon}_T. \tag{EC.51}$$

Combining (EC.49) with (EC.50), we see that for T sufficiently large,

$$\begin{aligned}\beta_T(c,d) &\leq \exp\left\{-(T-1)\bar{\alpha}_{T-1}(c)\left(G_d(\gamma(c),\bar{\beta}_{T-1}(c,d)) - \min_{d'\in\mathcal{D}_c\setminus\{d^*(c)\}} G_{d'}(\gamma(c),\bar{\beta}_{T-1}(c,d')) - 2\varepsilon_{T-1}\right)\right\}\\ &\leq \exp\{-(T-1)\bar{\alpha}_{T-1}(c)\left(G_d(\gamma(c),\bar{\beta}_{T-1}(c,d)) - \Gamma^*_{\gamma(c)} - \tilde{\varepsilon}_T - 2\varepsilon_{T-1}\right)\}\\ &\leq \exp\{-(T-1)\bar{\alpha}_{T-1}(c)\cdot\delta/2\},\end{aligned} \tag{EC.52}$$

where the second inequality follows from (EC.51). This results in negligible conditional sampling effort

$$\begin{aligned}\bar{\beta}_T(c,d) &= \frac{\sum_{t=1}^T \alpha_t(c)\beta_t(c,d)}{\sum_{t=1}^T \alpha_t(c)}\\ &= \frac{\sum_{t=1}^{\tilde{T}-1} \alpha_t(c)\beta_t(c,d)}{\sum_{t=1}^T \alpha_t(c)} + \frac{\sum_{t=\tilde{T}}^T \alpha_t(c)\beta_t(c,d)}{\sum_{t=1}^T \alpha_t(c)}\\ &\leq \frac{\sum_{t=1}^{\tilde{T}-1} \alpha_t(c)\beta_t(c,d)}{\sum_{t=1}^T \alpha_t(c)} + \frac{\sum_{t=\tilde{T}}^T \alpha_t(c)\exp\{-(t-1)\bar{\alpha}_{t-1}(c)\cdot\delta/2\}}{\sum_{t=1}^T \alpha_t(c)},\end{aligned}$$

where $\tilde{T}$ is an arbitrary positive integer such that the aforementioned asymptotic bounds hold for $t \geq \tilde{T}$. The first term in the equality vanishes eventually, while the second term can be bounded using the Abel transformation

$$\begin{aligned}&\frac{\sum_{t=\tilde{T}}^T \alpha_t(c)\exp\{-(T-1)\bar{\alpha}_{T-1}(c)\cdot\delta/2\}}{\sum_{t=1}^T \alpha_t(c)}\\ =&\frac{\alpha_T(c)\exp\{-(T-1)\bar{\alpha}_{T-1}(c)\cdot\delta/2\}}{\sum_{t=1}^T \alpha_t(c)}\\ &+\frac{\sum_{t=\tilde{T}}^{T-1}\left(\sum_{\tau=\tilde{T}}^t \alpha_\tau(c)\right)\left(\exp\{-(t-1)\bar{\alpha}_{t-1}(c)\cdot\delta/2\} - \exp\{-t\bar{\alpha}_t(c)\cdot\delta/2\}\right)}{\sum_{t=1}^T \alpha_t(c)}\\ \leq&\exp\{-(T-1)\bar{\alpha}_{T-1}(c)\cdot\delta/2\} + \sum_{t=\tilde{T}}^{T-1}\left(\exp\{-(t-1)\bar{\alpha}_{t-1}(c)\cdot\delta/2\} - \exp\{-t\bar{\alpha}_t(c)\cdot\delta/2\}\right)\\ =&\exp\{-(\tilde{T}-1)\bar{\alpha}_{\tilde{T}-1}(c)\cdot\delta/2\}.\end{aligned}$$

Therefore, we see that $\limsup_{T\to\infty}\bar{\beta}_T(c,d) \leq \exp\{-(\tilde{T}-1)\bar{\alpha}_{\tilde{T}-1}(c)\cdot\delta/2\}$. Note that $\tilde{T}$ can be chosen arbitrarily large. Sending $\tilde{T}\to\infty$, we see that $\lim_{T\to\infty}\bar{\beta}_T(c,d) = 0$ since $T\bar{\alpha}_T(c)\to\infty$ by Theorem 2. This results in

$$\lim_{T\to\infty} G_d(\bar{\beta}_T(c,d^*(c)),\bar{\beta}_T(c,d)) = G_d(\gamma(c),0) = 0,$$

which contradicts the supposition that $\liminf_{T\to\infty}\left(G_d(\gamma(c),\bar{\beta}_T(c,d)) - z_c\right)$ is positive.

Then we can show that $\limsup_{T\to\infty}\left(G_d(\gamma(c),\bar{\beta}_T(c,d)) - z_c\right)$ is non-positive. Again we argue by contradiction. Suppose otherwise that the limit superior is $\delta > 0$. This can only occur when there is an increasing sequence of time steps $T_1 < T_2 < \ldots < T_n < \ldots$, such that for any integer $k \geq 1$, $G_d(\gamma(c),\bar{\beta}_{T_{2k-1}}(c,d)) - \Gamma^*_{\gamma(c)} \leq \delta/2$, $G_d(\gamma(c),\bar{\beta}_{T_{2k}}(c,d)) - \Gamma^*_{\gamma(c)} \geq \delta$, and $\delta/2 < G_d(\gamma(c),\bar{\beta}_T(c,d)) - \Gamma^*_{\gamma(c)} < \delta$ for any $T_{2k-1} < T < T_{2k}$, since the corresponding limit superior is non-positive. Note that by definition,

$$\left|\bar{\beta}_{T+1} - \bar{\beta}_T\right| = \left|\frac{\sum_{t=1}^{T+1}\alpha_t(c)\beta_t(c,d)}{\sum_{t=1}^{T+1}\alpha_t(c)} - \frac{\sum_{t=1}^{T}\alpha_t(c)\beta_t(c,d)}{\sum_{t=1}^{T}\alpha_t(c)}\right| \leq \frac{1}{\sum_{t=1}^T\alpha_t(c)} \to 0, \tag{EC.53}$$

hence $|G_d(\gamma(c),\bar\beta_{T+1}(c,d))-G_d(\gamma(c),\bar\beta_T(c,d))|\to 0$. As a result, for sufficiently large k, we have $G_d(\gamma(c),\bar\beta_{T_{2k-1}}(c,d))-\Gamma^*_{\gamma(c)}>\delta/4$. Since $G_d(\cdot,\cdot)$ is uniformly continuous, there exists $\tilde\delta>0$ such that $|x-y|<\tilde\delta$ implies $|G_d(\gamma(c),x)-G_d(\gamma(c),y)|<\delta/2$. This in turn implies that

$$\bar\beta_{T_{2k}}(c,d)-\bar\beta_{T_{2k-1}}(c,d)\ge\tilde\delta.$$

Therefore, we have

$$\bar\beta_{T_{2k}}(c,d)=\frac{\sum_{t=1}^{T_{2k}}\alpha_t(c)\beta_t(c,d)}{\sum_{t=1}^{T_{2k}}\alpha_t(c)}\ge\bar\beta_{T_{2k-1}}(c,d)+\tilde\delta.$$

Rearranging the terms, we see that

$$\sum_{t=T_{2k-1}+1}^{T_{2k}}\alpha_t(c)\beta_t(c,d)\ge\sum_{t=1}^{T_{2k}}\alpha_t(c)\cdot(\bar\beta_{T_{2k-1}}(c,d)+\tilde\delta)-\sum_{t=1}^{T_{2k-1}}\alpha_t(c)\cdot\bar\beta_{T_{2k-1}}(c,d)\ge\sum_{t=T_{2k-1}+1}^{T_{2k}}\alpha_t(c)\cdot\tilde\delta.$$

However, similar to (EC.52), for $T_{2k-1}\le T\le T_{2k}$ and k sufficiently large,

$$\beta_T(c,d)\le\exp\{-(T-1)\bar\alpha_{T-1}(c)\cdot\delta/8\}\to 0,$$

leading to a contradiction.

Step 2.2. We have now established (EC.48). In the following, we will show that with probability 1,

$$\liminf_{T\to\infty}G_d(\gamma(c),\bar\beta_T(c,d))-\Gamma^*_{\gamma(c)}\ge 0,\quad\forall d\in\mathcal{D}_c\backslash\{d^*(c)\}.\tag{EC.54}$$

We start by showing that

$$\limsup_{T\to\infty}\min_{d'\in\mathcal{D}_c\backslash\{d^*(c)\}}G_{d'}(\gamma(c),\bar\beta_T(c,d'))-\Gamma^*_{\gamma(c)}\ge 0.\tag{EC.55}$$

Fix a $c\in\mathcal{C}$. Note that (EC.48) still holds if z_c is replaced with $\tilde z_c$, where

$$\tilde z_c:=\limsup_{T\to\infty}\min_{d'\in\mathcal{D}_c\backslash\{d^*(c)\}}G_{d'}(\gamma(c),\bar\beta_T(c,d'));$$

in other words,

$$\limsup_{T\to\infty}G_d(\gamma(c),\bar\beta_T(c,d))-\tilde z_c\le 0,\quad\forall d\in\mathcal{D}_c,\tag{EC.56}$$

and the proof is a simple repetition for (EC.48) in Step 2.1, which is omitted to avoid redundancy. Let $\beta^*_{\boldsymbol\gamma}:=(\beta^*_{\boldsymbol\gamma}(c,d))$ be an optimal solution to (11) that satisfies optimal condition (10). If $\tilde z_c\le\Gamma^*_{\gamma(c)}-\delta$ for some $\delta>0$, then

$$\limsup_{T\to\infty}G_d(\gamma(c),\bar\beta_T(c,d))\le\Gamma^*_{\gamma(c)}-\delta=G_d(\gamma(c),\beta^*_{\boldsymbol\gamma}(c,d))-\delta,\quad\forall d\in\mathcal{D}_c.$$

Since $G_d(\cdot,\cdot)$ is monotonically increasing and uniformly continuous, there exists a positive number $\varepsilon>0$ such that $\bar\beta_T(c,d)<\beta^*_{\boldsymbol\gamma}(c,d)-\varepsilon$ for each $d\in\mathcal{D}_c\backslash\{d^*(c)\}$. However,

$$\begin{aligned}1-\gamma(c)-(|\mathcal{D}_c|-1)\varepsilon&>\sum_{d\in\mathcal{D}_c\backslash\{d^*(c)\}}\bar\beta_T(c,d)\\&=1-\bar\beta_T(c,d^*(c))\to 1-\gamma(c),\end{aligned}$$

leading to a contradiction. Therefore, we have $\tilde z_c\ge\Gamma^*_{\gamma(c)}$, i.e., inequality (EC.55) holds.

To complete the proof, we leverage (EC.55) to show a stronger inequality, namely

$$\liminf_{T\to\infty}\min_{d'\in\mathcal{D}_c\backslash\{d^*(c)\}}G_{d'}(\gamma(c),\bar\beta_T(c,d'))-\Gamma^*_{\gamma(c)}\ge 0,$$

and thus (EC.54) follows immediately.

Again, let us argue by contradiction, and assume that the limit inferior is less than $-\delta$ for some $\delta > 0$. Then by (EC.55), there exist an increasing sequence of time steps $T_1 < T_2 < \ldots < T_n < \ldots$, such that for any integer $k \geq 1$, $\min_{d' \in \mathcal{D}_c \setminus \{d^*(c)\}} G_{d'}(\gamma(c), \bar{\beta}_{T_{2k-1}}(c,d')) - \Gamma^*_{\gamma(c)} \geq -\delta/2$, and $\min_{d' \in \mathcal{D}_c \setminus \{d^*(c)\}} G_{d'}(\gamma(c), \bar{\beta}_{T_{2k}}(c,d')) - \Gamma^*_{\gamma(c)} \leq -\delta$. Furthermore, $-\delta < \min_{d' \in \mathcal{D}_c \setminus \{d^*(c)\}} G_{d'}(\gamma(c), \bar{\beta}_{T}(c,d')) - \Gamma^*_{\gamma(c)} < -\delta/2$ for any $T_{2k-1} < T < T_{2k}$. Let d_k denote $\arg\min_{d' \in \mathcal{D}_c \setminus \{d^*(c)\}} G_{d'}(\gamma(c), \bar{\beta}_{T_{2k}}(c,d'))$. Then we see that

$$G_{d_k}(\gamma_c, \bar{\beta}_{T_{2k-1}}(c,d_k)) - G_{d_k}(\gamma_c, \bar{\beta}_{T_{2k}}(c,d_k)) > \delta/2.$$

Recall that $G_d(\cdot,\cdot)$ is universally uniformly continuous for all $d \in \mathcal{D}$ and monotonically increasing in both arguments. There exists a positive number $\varepsilon > 0$ such that for all $k \geq 1$, $\bar{\beta}_{T_{2k-1}}(c,d_k) - \bar{\beta}_{T_{2k}}(c,d_k) > \varepsilon$. By Lemma 1, for sufficiently large T and T',

$$\begin{aligned}
&\left| \sum_{d' \in \mathcal{D}_c \setminus \{d^*(c)\}} \bar{\beta}_T(c,d') - \sum_{d' \in \mathcal{D}_c \setminus \{d^*(c)\}} \bar{\beta}_{T'}(c,d') \right| \\
\leq &\left| \sum_{d' \in \mathcal{D}_c \setminus \{d^*(c)\}} \bar{\beta}_T(c,d') - (1-\gamma(c)) \right| + \left| \sum_{d' \in \mathcal{D}_c \setminus \{d^*(c)\}} \bar{\beta}_{T'}(c,d') - (1-\gamma(c)) \right| \\
\leq &\varepsilon/2.
\end{aligned}$$

Hence there exists $d'_k \neq d_k$ such that $\bar{\beta}_{T_{2k-1}}(c,d'_k) - \bar{\beta}_{T_{2k-1}}(c,d'_k) < -\varepsilon/(2|\mathcal{D}_c|)$. By (EC.53), for sufficiently large k, there exists a time step $\tilde{T}_k$ with $T_{2k-1} \leq \tilde{T}_k < T_{2k}$, such that $\forall \tilde{T}_k \leq T < T_{2k}$,

$$G_{d'_k}(\gamma_c, \bar{\beta}_T(c,d'_k)) - \Gamma^*_{\gamma(c)} \geq -\frac{2}{3}\delta,$$

and

$$G_{d_k}(\gamma_c, \bar{\beta}_T(c,d_k)) - \Gamma^*_{\gamma(c)} \leq -\frac{5}{6}\delta.$$

This leads to $\beta_T(c,d'_k) \leq \exp\{-(T-1)\bar{\alpha}_{T-1} \cdot \delta/12\}$ for sufficiently large k, and $\tilde{T}_k < T < T_{2k}$. However, it contradicts $\bar{\beta}_{T_{2k-1}}(c,d'_k) - \bar{\beta}_{T_{2k}}(c,d'_k) < -\varepsilon/(2|\mathcal{D}_c|)$, which leads to

$$\sum_{t=\tilde{T}_k+1}^{T_{2k}} \alpha_t(c)\beta_t(c,d'_k) \geq \sum_{t=\tilde{T}_k+1}^{T_{2k}} \left(\alpha_t(c) \left(\bar{\beta}_T(c,d'_k) + \frac{\varepsilon}{2|\mathcal{D}_c|} \right) \right).$$

To partially conclude, we have

$$\lim_{T \to \infty} G_d(\gamma(c), \bar{\beta}_T(c,d)) = z_c = \Gamma^*_{\gamma(c)}, \quad \forall d \in \mathcal{D}_c \setminus \{d^*(c)\}.$$

Equivalently, the conditional sampling effort leads to the optimal value of rate function almost surely in each context. Moreover, it follows from the universally uniform continuity of $G_d(\cdot,\cdot)$'s that $\lim_{T\to\infty} G_d(\bar{\beta}_T(c,d^*(c)), \bar{\beta}_T(c,d)) - G_d(\gamma(c), \bar{\beta}_T(c,d)) = 0$, and therefore,

$$\lim_{T \to \infty} G_d(\bar{\beta}_T(c,d^*(c)), \bar{\beta}_T(c,d)) = z_c = \Gamma^*_{\gamma(c)}, \quad \forall d \in \mathcal{D}_c \setminus \{d^*(c)\}.$$

Step 3: Global balance of sampling ratios to contexts. We complete the proof by showing that $\lim_{T\to\infty} \bar{\alpha}_T(c) = \alpha^*_{\boldsymbol{\gamma}}(c)$ for any fixed context $c \in \mathcal{C}$. According to Corollary 1,

$$1 - \prod_{c' \in \mathcal{C}} \pi^{(T-1)}_{d^*(c')} = \Pi_{T-1} \left(\bigcup_{c' \in \mathcal{C}} \bigcup_{d' \in \mathcal{D}_{c'} \setminus \{d^*(c')\}} \mathbf{\Theta}_{(c',d')} \right)$$

$$
\begin{aligned}
&\le \exp\left\{-(T-1)\left(\min_{c'\in\mathcal{C}}\bar{\alpha}_{T-1}(c')\min_{d'\in\mathcal{D}_{c'}\setminus\{d^*(c')\}}G_{d'}(\bar{\beta}_{T-1}(c',d^*(c')),\bar{\beta}_{T-1}(c',d'))-\frac{\varepsilon_{T-1}}{2}\right)\right\}\\
&\le \exp\left\{-(T-1)\left(\min_{c'\in\mathcal{C}}\bar{\alpha}_{T-1}(c')\Gamma^*_{\gamma(c')}-\varepsilon_{T-1}\right)\right\},
\end{aligned}
$$

where $\{\varepsilon_T\}_{T\ge 1}$ is a vanishing sequence of positive numbers, and the last inequality follows from Step 2. Similarly, we also have

$$
1-\prod_{c'\in\mathcal{C}}\pi^{(T-1)}_{d^*(c')}\ge \exp\left\{-(T-1)\left(\min_{c'\in\mathcal{C}}\bar{\alpha}_{T-1}(c')\Gamma^*_{\gamma(c')}+\varepsilon_{T-1}\right)\right\}. \tag{EC.57}
$$

Furthermore, it also follows by taking $\tilde{\boldsymbol{\Theta}}=\boldsymbol{\Theta}^{(c,d^*(c))}$ in Theorem 1 from the conclusion of Step 2 that

$$
\exp\left\{-(T-1)\left(\bar{\alpha}_{T-1}(c)\Gamma^*_{\gamma(c)}+\varepsilon_{T-1}\right)\right\}\le 1-\pi^{(T-1)}_{d^*(c)}\le \exp\left\{-(T-1)\left(\bar{\alpha}_{t-1}(c)\Gamma^*_{\gamma(c)}-\varepsilon_{T-1}\right)\right\}. \tag{EC.58}
$$

We first show that $\liminf_{T\to\infty}\bar{\alpha}_T(c)\le\alpha^*_{\boldsymbol{\gamma}}(c)$. Assume, for the sake of contradiction, that there exist some $\delta>0$, $\bar{\alpha}_T(c)>\alpha^*_{\boldsymbol{\gamma}}(c)(1+\delta)$ for any sufficiently large T. Note also that $\Gamma^*_{\boldsymbol{\gamma}}\ge\min_{c'\in\mathcal{C}}\bar{\alpha}_{T-1}(c')\Gamma^*_{\gamma(c')}$ by definition (15). Hence

$$
\bar{\alpha}_{T-1}(c)\Gamma^*_{\gamma(c)}-\min_{c'\in\mathcal{C}}\bar{\alpha}_{T-1}(c')\Gamma^*_{\gamma(c')}>(1+\delta)\alpha^*_{\boldsymbol{\gamma}}(c)\Gamma^*_{\gamma(c)}-\Gamma^*_{\boldsymbol{\gamma}}=\delta\cdot\Gamma^*_{\boldsymbol{\gamma}},
$$

where the last equality follows from $\alpha^*_{\boldsymbol{\gamma}}(c)\Gamma^*_{\gamma(c)}=\Gamma^*_{\boldsymbol{\gamma}}$ by definition (14).

Combining Lemma 2 with inequalities (EC.57) and (EC.58), we obtain $\alpha_T(c)\le\exp\{-(T-1)(\delta\Gamma^*_{\boldsymbol{\gamma}}-2\varepsilon_{T-1})\}\le\exp\{-(T-1)\cdot\delta\Gamma^*_{\boldsymbol{\gamma}}/2\}$, which implies that $\lim_{T\to\infty}\bar{\alpha}_T(c)=0$, leading to a contradiction to the assumption that $\bar{\alpha}_T(c)>\alpha^*_{\boldsymbol{\gamma}}(c)(1+\delta)$.

We further show that $\limsup_{T\to\infty}\bar{\alpha}_T(c)\le\alpha^*_{\boldsymbol{\gamma}}(c)$. Otherwise, there exist some $\delta>0$ and an increasing sequence of time steps $T_1<T_2<\ldots<T_n<\ldots$, such that for any integer $k\ge 1$, $\bar{\alpha}_{T_{2k-1}}(c)\le\alpha^*_{\boldsymbol{\gamma}}(c)(1+\delta/2)$, $\bar{\alpha}_{T_{2k}}(c)\ge\alpha^*_{\boldsymbol{\gamma}}(c)(1+\delta)$, and for all $T_{2k-1}<T<T_{2k}$, $\alpha^*_{\boldsymbol{\gamma}}(c)(1+\delta/2)<\bar{\alpha}_{T_{2k}}(c)<\alpha^*_{\boldsymbol{\gamma}}(c)(1+\delta)$. Since $|\bar{\alpha}_{T+1}(c)-\bar{\alpha}_T(c)|\le 1/(T+1)\to 0$, for any sufficiently large k, $\bar{\alpha}_{T_{2k-1}}(c)\ge\alpha^*_{\boldsymbol{\gamma}}(c)(1+\delta/4)$. As a result, $\alpha_T(c)\le\exp\{-(T-1)\cdot\delta\Gamma^*_{\boldsymbol{\gamma}}/8\}$ for $T_{2k-1}<T<T_{2k}$. According to (EC.57) and (EC.58),

$$
\begin{aligned}
\delta\cdot\Gamma^*_{\boldsymbol{\gamma}}/2<\bar{\alpha}_{T_{2k}}(c)-\bar{\alpha}_{T_{2k-1}}(c)&=\frac{1}{T_{2k}}\sum_{t=1}^{T_{2k}}\alpha_t(c)-\frac{1}{T_{2k-1}}\sum_{t=1}^{T_{2k-1}}\alpha_t(c)\\
&=\frac{T_{2k}-T_{2k-1}}{T_{2k}}\left(\frac{1}{T_{2k}-T_{2k-1}}\sum_{t=T_{2k-1}+1}^{T_{2k}}\alpha_t(c)-\frac{1}{T_{2k-1}}\sum_{t=1}^{T_{2k-1}}\alpha_t(c)\right)\\
&\le\frac{T_{2k}-T_{2k-1}}{T_{2k}}\exp\{-(T-1)\cdot\delta\Gamma^*_{\boldsymbol{\gamma}}/8\}\to 0,
\end{aligned}
$$

leading to a contradiction to the assumption that $\delta>0$.

We come to the conclusion that $\limsup_{T\to\infty}\bar{\alpha}_T(c)\le\alpha^*(c)$. Note that by definition, $\sum_{c\in\mathcal{C}}\bar{\alpha}_T(c)=1$, there must be $\lim_{T\to\infty}\bar{\alpha}_T(c)=\alpha^*(c)$. □

Remark EC.1. The proof can, to a much extent, be simplified if we assume uniqueness of the optimal allocation and strict monotonicity of rate functions. In such cases, Step 2.1 implies that $\limsup_{T\to\infty}\bar{\beta}_T(c,d)\le\beta^*(c,d)$, and immediately we have $\lim_{T\to\infty}\bar{\beta}_T(c,d)=\beta^*(c,d)$ since $\sum_{d\in\mathcal{D}_c}\bar{\beta}_T(c,d)=\sum_{d\in\mathcal{D}_c}\beta^*(c,d)=1$ for all $T\ge 1$ and $c\in\mathcal{C}$. Thus, Step 2.2 can be omitted, and this reduces to a similar argument to Russo [45].

Then, we have a direct result for Algorithm 2 and Algorithm 4 with a tuning hyperparameter.

Corollary EC.3 **(Asymptotic Optimality)**. *For TTTS-C with tuning hyperparameter in Algorithm 2 and Algorithm 4, we have*

$$\lim_{T\to\infty} -\frac{1}{T}\ln \Pi_T\left(\bigcup_{c\in\mathcal{C}}\bigcup_{d\in\mathcal{D}\setminus\{d^*(c)\}}\boldsymbol{\Theta}_{(c,d)}\right)=\Gamma^*.$$

For any sequential sampling policy,

$$\limsup_{T\to\infty} -\frac{1}{T}\ln \Pi_T\left(\bigcup_{c\in\mathcal{C}}\bigcup_{d\in\mathcal{D}\setminus\{d^*(c)\}}\boldsymbol{\Theta}_{(c,d)}\right)\le\Gamma^*.$$

Proof. Let $\boldsymbol{\gamma}_t:=(\gamma_t(c))$ be the hyperparameter used at time t. Let $\hat{\boldsymbol{\theta}}_t$ be the consistent estimate of $\boldsymbol{\theta}^*$. According to Theorem EC.1, each design $d\in\mathcal{D}$ receives infinite simulation samples ultimately with probability 1. Therefore, $\hat{\boldsymbol{\theta}}_t\to\boldsymbol{\theta}^*$ with probability 1. It follows from the assumption $\mu^*_{d^*(c)}>\max_{d\in\mathcal{D}_c}\mu^*_d$, $\forall c\in\mathcal{C}$ that there exists a random integer $\tilde{T}$ such that for all $t\ge\tilde{T}$, $c\in\mathcal{C}$ and $d\in\mathcal{D}_c\setminus\{d^*(c)\}$, $(\hat{\mu}_t)_{d^*(c)}>\frac{1}{2}\mu^*_{d^*(c)}+\frac{1}{2}\max_{d\in\mathcal{D}_c}\mu^*_d>(\hat{\mu}_t)_d$. Therefore, the plug-in estimation $\hat{\boldsymbol{d}}_t:=(\hat{d}_t(c))$ of $\boldsymbol{d}^*$ is consistent. It turns out that $\lim_{t\to\infty}\Gamma^*_{\boldsymbol{\gamma}_t}=\Gamma^*_{\boldsymbol{\gamma}^*}=\Gamma^*$, according to the envelop theorem [see 33, Corollary 3], since $\boldsymbol{\gamma}_t(c)$ equates to the estimated optimal sampling ratio of contextual best designs solved by maximizing (8) with parameters substituted by their estimates. Now, what remains is to simply repeat the proof of Corollary 3, which completes the proof of the first statement.

The second statement is a direct consequence of Corollary 1 and the definition of Γ^*. ☐

Appendix E: Extension to top-m_c selection. Now, we can extend the analysis for the large deviations rate for the generic top-m_c selection problem. To this end, the sampling policy can be expressed explicitly and their bounds are derived as in Lemmas 1 and 2.

Lemma EC.4. *For any $c\in\mathcal{C}$ and $d\in\mathcal{D}_c$, the conditional sampling ratio of TTTS-C with hyperparameter $\boldsymbol{\gamma}$ satisfies*

1. *for $d\in\mathcal{P}_c$, we have*

$$\lim_{t\to\infty}\sum_{d\in\mathcal{P}_c}\beta_t(c,d)=\gamma(c),$$

and with probability 1, for sufficiently large t,

$$\frac{\gamma(c)}{|\mathcal{D}_c|^2}\frac{\sum_{\tilde{P}_c:d\notin\tilde{P}_c}\pi^{(t-1)}_{\tilde{P}_c}}{1-\pi^{(t-1)}_{\mathcal{P}_c}}\le\beta_t(c,d)\le 2\frac{\sum_{\tilde{P}_c:d\notin\tilde{P}_c}\pi^{(t-1)}_{\tilde{P}_c}}{1-\pi^{(t-1)}_{\mathcal{P}_c}},$$

2. *for $d'\in\mathcal{U}_c$, we have*

$$\lim_{t\to\infty}\sum_{d'\in\mathcal{U}_c}\beta_t(c,d')=1-\gamma(c),$$

and with probability 1, for sufficiently large t,

$$\frac{1-\gamma(c)}{|\mathcal{D}_c|^2}\frac{\sum_{\tilde{P}_c:d'\in\tilde{P}_c}\pi^{(t-1)}_{\tilde{P}_c}}{1-\pi^{(t-1)}_{\mathcal{P}_c}}\le\beta_t(c,d)\le 2\frac{\sum_{\tilde{P}_c:d'\in\tilde{P}_c}\pi^{(t-1)}_{\tilde{P}_c}}{1-\pi^{(t-1)}_{\mathcal{P}_c}}.$$

Moreover, with probability 1, for sufficiently large t,

$$\frac{1}{2}\frac{1-\pi^{(t-1)}_{\mathcal{P}_c}}{1-\prod_{c'\in\mathcal{C}}\pi^{(t-1)}_{\mathcal{P}_{c'}}}\le\alpha_t(c)\le 2\frac{1-\pi^{(t-1)}_{\mathcal{P}_c}}{1-\prod_{c'\in\mathcal{C}}\pi^{(t-1)}_{\mathcal{P}_{c'}}}.$$

Proof. We only prove for the two limits. For $c\in\mathcal{C}$ and a subset $P_c\subseteq\mathcal{D}_c$ with $|P_c|=m_c$, let $\pi_{P_c}^{(t-1)}:=\Pi_{t-1}(\mathbf{\Theta}^{(c,P_c)})$ denote the posterior probability of subset $P_c\subseteq\mathcal{D}_c$ being equal to $\mathcal{P}_c$. Then similar to Lemma EC.3, we have the following expression for the sampling policy under TTTS-C for selecting top-m_c designs. For any fixed $c\in\mathcal{C}$ and $d\in\mathcal{D}_c$,

$$\begin{aligned}\psi_t(d)=&\sum_{P:d\in P_c}\prod_{c'\in\mathcal{C}}\pi_{P_{c'}}^{(t-1)}\sum_{\Delta\ni c}\frac{\prod_{c'\in\Delta}\left(1-\pi_{P_{c'}}^{(t-1)}\right)\prod_{c'\notin\Delta}\pi_{P_{c'}}^{(t-1)}}{1-\prod_{c'\in\mathcal{C}}\pi_{P_{c'}}^{(t-1)}}\cdot\frac{\sum_{\tilde P_c:d\notin\tilde P_c}\pi_{\tilde P_c}^{(t-1)}/|P_c-\tilde P_c|}{1-\pi_{P_c}^{(t-1)}}\frac{\gamma(c)}{|\Delta|}\\&+\sum_{P:d\notin P_c}\prod_{c'\in\mathcal{C}}\pi_{P_{c'}}^{(t-1)}\sum_{\Delta\ni c}\frac{\prod_{c'\in\Delta}\left(1-\pi_{P_{c'}}^{(t-1)}\right)\prod_{c'\notin\Delta}\pi_{P_{c'}}^{(t-1)}}{1-\prod_{c'\in\mathcal{C}}\pi_{P_{c'}}^{(t-1)}}\cdot\frac{\sum_{\tilde P_c:d\in\tilde P_c}\pi_{\tilde P_c}^{(t-1)}/|\tilde P_c-P_c|}{1-\pi_{P_c}^{(t-1)}}\frac{1-\gamma(c)}{|\Delta|},\end{aligned}\tag{EC.59}$$

and

$$\alpha_t(c)=\sum_P\frac{\prod_{c'\in\mathcal{C}}\pi_{P_{c'}}^{(t-1)}}{1-\prod_{c'\in\mathcal{C}}\pi_{P_{c'}}^{(t-1)}}\sum_{\Delta\ni c}\prod_{c'\in\Delta}\left(1-\pi_{P_{c'}}^{(t-1)}\right)\prod_{c'\notin\Delta}\pi_{P_{c'}}^{(t-1)}\frac{1}{|\Delta|},\tag{EC.60}$$

where the summations $\sum_P$, $\sum_{P:d\in P_c}$, and $\sum_{P:d\notin P_c}$ are taken over all combinations (that satisfy the corresponding condition) of subsets $P_{c'}\subseteq\mathcal{D}_{c'}$ with $|P_{c'}|=m_{c'}$ for context c' ranging over $\mathcal{C}$, while the summation $\sum_{\tilde P_c:d\in\tilde P_c}$ is taken over subsets $\tilde P_c\subseteq\mathcal{D}_c$ with $|\tilde P_c|=m_c$ that satisfies $d\in\tilde P_c$, for the fixed context c and design d.

Note that for $d\in\mathcal{P}_c$, the second summation in Equation (EC.59) converges to 0 due to Theorem EC.1. Therefore, we have

$$\begin{aligned}&\lim_{t\to\infty}\sum_{d\in\mathcal{P}_c}\psi_t(d)\\=&\lim_{t\to\infty}\sum_{d\in\mathcal{P}_c}\sum_{P:d\in P_c}\prod_{c'\in\mathcal{C}}\pi_{P_{c'}}^{(t-1)}\sum_{\Delta\ni c}\frac{\prod_{c'\in\Delta}\left(1-\pi_{P_{c'}}^{(t-1)}\right)\prod_{c'\notin\Delta}\pi_{P_{c'}}^{(t-1)}}{1-\prod_{c'\in\mathcal{C}}\pi_{P_{c'}}^{(t-1)}}\\&\qquad\cdot\frac{\sum_{\tilde P_c:d\notin\tilde P_c}\pi_{\tilde P_c}^{(t-1)}/|P_c-\tilde P_c|}{1-\pi_{P_c}^{(t-1)}}\frac{\gamma(c)}{|\Delta|}\\=&\lim_{t\to\infty}\sum_P\prod_{c'\in\mathcal{C}}\pi_{P_{c'}}^{(t-1)}\sum_{\Delta\ni c}\frac{\prod_{c'\in\Delta}\left(1-\pi_{(c',P_{c'})}^{(t-1)}\right)\prod_{c'\notin\Delta}\pi_{P_{c'}}^{(t-1)}}{1-\prod_{c'\in\mathcal{C}}\pi_{P_{c'}}^{(t-1)}}\\&\qquad\cdot\sum_{\tilde P_c:\tilde P_c\neq P_c}\frac{\pi_{\tilde P_c}^{(t-1)}}{1-\pi_{P_c}^{(t-1)}}\frac{|(P_c-\tilde P_c)\bigcap\mathcal{P}_c|}{|P_c-\tilde P_c|}\frac{\gamma(c)}{|\Delta|}.\end{aligned}$$

An analogous argument as in the proof for Lemma EC.3 yields that the summand with $P_c=\mathcal{P}_c$ dominates the summation, i.e.,

$$\begin{aligned}&\lim_{t\to\infty}\sum_{d\in\mathcal{P}_c}\beta_t(c,d)=\lim_{t\to\infty}\sum_{d\in\mathcal{P}_c}\psi_t(d)/\alpha_t(c)\\=&\lim_{t\to\infty}\frac{\prod\limits_{c'\in\mathcal{C}}\pi_{\mathcal{P}_{c'}}^{(t-1)}\sum\limits_{\Delta\ni c}\frac{\prod\limits_{c'\in\Delta}\left(1-\pi_{\mathcal{P}_{c'}}^{(t-1)}\right)\prod\limits_{c'\notin\Delta}\pi_{\mathcal{P}_{c'}}^{(t-1)}}{1-\prod\limits_{c'\in\mathcal{C}}\pi_{\mathcal{P}_{c'}}^{(t-1)}}\cdot\sum\limits_{\tilde P_c:\tilde P_c\neq\mathcal{P}_c}\frac{\pi_{\tilde P_c}^{(t-1)}}{1-\pi_{\mathcal{P}_c}^{(t-1)}}\frac{|(\mathcal{P}_c-\tilde P_c)\bigcap\mathcal{P}_c|}{|\mathcal{P}_c-\tilde P_c|}\frac{\gamma(c)}{|\Delta|}}{\prod\limits_{c'\in\mathcal{C}}\pi_{\mathcal{P}_{c'}}^{(t-1)}\sum\limits_{\Delta\ni c}\frac{\prod\limits_{c'\in\Delta}\left(1-\pi_{\mathcal{P}_{c'}}^{(t-1)}\right)\prod\limits_{c'\notin\Delta}\pi_{\mathcal{P}_{c'}}^{(t-1)}}{1-\prod\limits_{c'\in\mathcal{C}}\pi_{\mathcal{P}_{c'}}^{(t-1)}}\cdot\frac{1}{|\Delta|}}\\=&\gamma(c),\end{aligned}$$

where the last equation follows from $(\mathcal{P}_c - \tilde{P}_c)\bigcap \mathcal{P}_c = \mathcal{P}_c - \tilde{P}_c$ and $\sum_{\tilde{P}_c:\tilde{P}_c\neq\mathcal{P}_c} \pi_{\tilde{P}_c}^{(t-1)}/\left(1-\pi_{\mathcal{P}_c}^{(t-1)}\right)=1$. It follows from $\sum_{d\in\mathcal{D}_c}\beta_t(c,d)=1$, $\forall c\in\mathcal{C}$ that $\sum_{d'\in\mathcal{U}_c}\beta_t(c,d')\to 1-\gamma(c)$.

As for the inequalities, the proof resembles those for Lemma 1 and Lemma 2, which is thus omitted. $\square$

Proof for Theorem 4. The proof also breaks down into three steps. The first step is to establish the consistency of TTTS-C, which is a direct consequence of Theorem EC.1. Therefore, for all $c\in\mathcal{C}$, we have $\pi_{\mathcal{P}_c}^{(t-1)}\to 1$. The second step and the third step, respectively, are to show that for all $c\in\mathcal{C}$, $d\in\mathcal{P}_c$, $d'\in\mathcal{U}_c$,

$$\lim_{T\to\infty}\min_{\tilde{d}\in\mathcal{P}_c} G_{\tilde{d},d'}(\bar{\beta}_T(c,\tilde{d}),\bar{\beta}_T(c,d')) - z_{T,c} = \lim_{T\to\infty}\min_{\tilde{d}'\in\mathcal{U}_c} G_{d,\tilde{d}'}(\bar{\beta}_T(c,d),\bar{\beta}_T(c,\tilde{d}')) - z_{T,c} = 0,$$

where $z_{T,c} := \min_{\tilde{d}\in\mathcal{P}_c}\min_{\tilde{d}'\in\mathcal{U}_c} G_{\tilde{d},\tilde{d}'}(\bar{\beta}_T(c,\tilde{d}),\bar{\beta}_T(c,\tilde{d}'))$, and that for all $c\in\mathcal{C}$,

$$\lim_{T\to\infty}\bar{\alpha}_T(c)z_{T,c} - z_T = 0,$$

where $z_T := \min_{c'\in\mathcal{C}}\bar{\alpha}_T(c)z_{T,c}$. The argument repeats that of Theorem 2 with Lemma EC.4 in place of Lemma 1 and Lemma 2, which completes the proof. $\square$

Appendix F: Discussions on top-m_c selection. In this section, we further discuss the optimal sampling ratios for top-m_c selection problems. Recall that the optimal sampling ratio for each context is inversely proportional to the optimal rate attained by allocating samples within each context. Therefore, we only have to focus on the context-free case, where $|\mathcal{C}|=1$. Following this section, we leave out the subscript c for simplicity.

F.1. Uniqueness of optimal allocation. We first consider the uniqueness of optimal allocation problem associated with the context-free top-m_c selection:

$$\begin{aligned} \max_{\boldsymbol{\psi}} \quad & G(\boldsymbol{\psi}) \\ \text{s.t.} \quad & \sum_{d\in\mathcal{P}}\psi(d) + \sum_{d'\in\mathcal{U}}\psi(d) = 1, \\ & \boldsymbol{\psi}\geq 0, \end{aligned} \tag{EC.61}$$

where

$$G(\boldsymbol{\psi}) := \min_{(d,d')\in\mathcal{P}\times\mathcal{U}} G_{d,d'}(\psi(d),\psi(d')).$$

Proposition EC.1. *Suppose $G_{d,d'}(\cdot,\cdot)$'s are strictly concave as bivariate functions, then the solution to the optimization problem (EC.61) is unique.*

Recall that the optimal sampling ratios of (18) for each context is uniquely determined by being inversely proportional to $\Gamma^*_{\gamma^*(c)}$, and thus we only have to show the optimal conditional sampling ratio for each context in a context is unique. Then Proposition 5 is a direct consequence of Proposition EC.1.

The proof leverages the concave property of the objective $G(\boldsymbol{\psi})$, although it may not be strictly concave as a multivariate function. The idea is to iteratively eliminate all potentially variable coordinates of the optima by utilizing the strict concavity of rate functions.

Proof of Proposition EC.1. We argue by contradiction. Assume that $\boldsymbol{\psi}^{(0)}$ and $\boldsymbol{\psi}^{(1)}$ are two distinct optimal solutions to problem (EC.61). Then $\Gamma^* = G(\boldsymbol{\psi}^{(0)}) = G(\boldsymbol{\psi}^{(1)})$ by definition. With a slight abuse of notations, define by

$$\boldsymbol{\psi}^{(r)} := r\boldsymbol{\psi}^{(1)} + (1-r)\boldsymbol{\psi}^{(0)}, \quad r\in[0,1],$$

the convex combinations of $\boldsymbol{\psi}^{(0)}$ and $\boldsymbol{\psi}^{(1)}$.

Observe that $\boldsymbol{\psi}^{(r)}$ maximizes the objective for $r \in [0,1]$. It simply follows from

$$\Gamma^* \geq G(\boldsymbol{\psi}^{(r)}) \geq rG(\boldsymbol{\psi}^{(1)}) + (1-r)G(\boldsymbol{\psi}^{(0)}) = \Gamma^*,$$

where the second inequality holds due to the fact that $G(\cdot)$, as a function minimizing over concave functions $G_{d,d'}(\cdot,\cdot)$, is concave in $\boldsymbol{\psi}$ as well.

Then we can inductively construct nest intervals $[r_{L,n}, r_{R,n}]$ along with a sequence of partitions of $\mathcal{P}\times\mathcal{U}$, i.e., $\{(A_n, B_n) : n \in \mathbb{N}\}$ such that $A_n \bigcup B_n$ is a disjoint union equal to $\mathcal{P}\times\mathcal{U}$. By construction, we will show that (a) A_n has cardinality n; (b) A_n consists of (d,d')-pairs such that the coordinates corresponding to d (and d') in $\boldsymbol{\psi}^{(0)}$ and $\boldsymbol{\psi}^{(1)}$ coincide; and (c) for any $n \geq 1$ and $r \in [r_{L,n}, r_{R,n}]$, there holds

$$\Gamma^* = \min_{(d,d')\in A_n} G_{d,d'}(\psi^{(r)}(d), \psi^{(r)}(d')).$$

A_n can been deemed as a desired set of indices whose corresponding coordinates in optimal solutions must coincide. On the other hand, B_n is a suspicious set whose elements correspond to coordinates in $\boldsymbol{\psi}^{(r)}$ that may vary. The induction procedure goes on until we can guarantee that $\boldsymbol{\psi}^{(0)} = \boldsymbol{\psi}^{(1)}$.

To be specific, let $r_{L,0} = 0$, $r_{R,0} = 1$, $P_0 = \emptyset$, and $Q_0 = \mathcal{P}\times\mathcal{U}$. We then provide the inductive step in constructing these quantities. Assume that the aforementioned quantities are well defined up to index $(n-1)$. We claim that at least one of the following case happens:

Case 1. $\psi^{(0)}(d) = \psi^{(1)}(d)$, and $\psi^{(0)}(d') = \psi^{(1)}(d')$, $\forall (d,d') \in B_{n-1}$;

Case 2. $\min_{(d,d')\in B_{n-1}} G_{d,d'}(\psi^{(r)}(d), \psi^{(r)}(d')) = \Gamma^*$, $\forall r \in (r_{L,n-1}, r_{R,n-1})$.

For $n = 1$, the claim is trivial since Case 2 is true. For $n \geq 2$, we argue by contradiction. Assume otherwise that neither of this two cases happens, that is, there exists an $(\tilde{d}, \tilde{d}')$-pair in B_{n-1} such that at least one of $\psi^{(0)}(\tilde{d}) \neq \psi^{(1)}(\tilde{d})$ and $\psi^{(0)}(\tilde{d}') \neq \psi^{(1)}(\tilde{d}')$ is true, and for some $r^0 \in (r_{L,n-1}, r_{R,n-1})$,

$$\min_{(d,d')\in B_{n-1}} G_{d,d'}(\psi^{(r^0)}(d), \psi^{(r^0)}(d')) > G(\boldsymbol{\psi}^{(r^0)}) = \Gamma^*.$$

Without loss of generality, we assume $\psi^{(0)}(\tilde{d}) < \psi^{(1)}(\tilde{d})$, and therefore $\psi^{(r)}(\tilde{d})$ is monotonically increasing in r.

Since $G_{d,d'}(\cdot,\cdot)$'s and $G(\cdot)$ are continuous functions, there exists $\varepsilon > 0$ such that for all $\tilde{\boldsymbol{\psi}}$ with $\|\tilde{\boldsymbol{\psi}} - \boldsymbol{\psi}^{(r^0)}\|_\infty \leq \varepsilon$,

$$\begin{aligned}&\sum_{(d,d')\in\mathcal{P}\times\mathcal{U}} \left| G_{d,d'}(\tilde{\psi}(d), \tilde{\psi}(d')) - G_{d,d'}(\psi^{(r^0)}(d), \psi^{(r^0)}(d')) \right| + \left| G(\tilde{\boldsymbol{\psi}}) - G(\boldsymbol{\psi}^{(r^0)}) \right| \\ &< \min_{(d,d')\in B_{n-1}} G_{d,d'}(\psi^{(r^0)}(d), \psi^{(r^0)}(d')) - \Gamma^*. \end{aligned} \tag{EC.62}$$

Let R_{n-1} be the set of indices in elements of A_{n-1}, that is,

$$R_{n-1} := \left\{\bar{d} : (\bar{d}, d') \in A_{n-1} \text{ for some } d' \in \mathcal{U} \text{ or } (d, \bar{d}) \in A_{n-1} \text{ for some } d \in \mathcal{P}\right\}.$$

Note that by (b), $d \notin R_{n-1}$, i.e., there exists no $\bar{d}$ such that $(d, \bar{d}) \in A_{n-1}$. Then one can choose $0 < \delta < \varepsilon$ to be small enough so that $\tilde{\boldsymbol{\psi}}$ is a feasible solution to (EC.61) given by

$$\tilde{\psi}(\bar{d}) = \begin{cases} \psi^{(r^0)}(d) - \delta, & \text{if } \bar{d} = d, \\ \psi^{(r^0)}(\bar{d}) + \delta/|R_{n-1}|, & \text{if } \bar{d} \in R_{n-1}, \\ \psi^{(r^0)}(\bar{d}), & \text{otherwise.} \end{cases} \tag{EC.63}$$

This solution yields a larger objective

$$G(\tilde{\boldsymbol{\psi}}) = \min_{(d,d')\in A_{n-1}} G_{d,d'}(\tilde{\psi}(d),\tilde{\psi}(d')) > \min_{(d,d')\in A_{n-1}} G_{d,d'}(\psi^{(r^0)}(d),\psi^{(r^0)}(d')) = \Gamma^*, \tag{EC.64}$$

which contradicts the definition of Γ^*. The first equation follows from triangular inequality and (EC.62), and the second inequality follows from the fact that $G_{d,d'}(\cdot,\cdot)$ is strictly increasing in both arguments.

Now, we have shown that at least one of cases must be true. In Case 1, we terminate the construction procedure. In Case 2, we define for each (d,d')-pair in B_{n-1} a subset of $[r_{L,n-1}, r_{R,n-1}]$ as

$$W^n_{d,d'} := \left\{ r\in[r_{L,n-1},r_{R,n-1}] : G_{d,d'}(\psi^{(r)}(d),\psi^{(r)})(d') = \min_{(i',j')\in B_{n-1}} G_{i'j'}(\psi^{(r)}_{i'},\psi^{(r)}_{j'}) = \Gamma^* \right\}.$$

Obviously $\bigcup_{(d,d')\in B_{n-1}} W^n_{d,d'} = [r_{L,n-1}, r_{R,n-1}]$, and for each $(d,d')\in B_{n-1}$, $W^n_{d,d'}$ is a closed interval since $G_{d,d'}(\psi^{(r)}_i,\psi^{(r)}_j)$ is continuous and concave in r. Then the Baire category theorem [see 34, Theorem 48.2] implies that there exists at least one pair $(d_n,d'_n)\in B_{n-1}$ such that $W^n_{d_n,d'_n}$ has non-empty interior. Now, we can define

$$\begin{aligned} &r_{L,n} := \inf W^n_{d_n,d'_n}, \quad r_{R,n} := \sup W^n_{d_n,d'_n}, \\ &A_n = A_{n-1}\bigcup\{(d_n,d'_n)\}, \quad B_n = B_{n-1}\backslash\{(d_n,d'_n)\}. \end{aligned}$$

Note that (a) and (c) hold naturally by construction. To justify (b), it suffices to show that $\psi^{(0)}(d_n) = \psi^{(1)}(d'_n)$. By the definition of $W^n_{d_n,d'_n}$, we see that $G_{d_n,d'_n}(\psi^{(r)}(d_n),\psi^{(r)}(d'_n)) = \Gamma^*$ for $r\in[r_{L,n}, r_{R,n}]$. Note that if $\psi^{(r^{L,n})}(d_n)\neq\psi^{(r^{R,n})}(d_n)$ or $\psi^{(r^{L,n})}(d'_n)\neq\psi^{(r^{R,n})}(d'_n)$, then

$$\begin{aligned} \Gamma^* &= G_{d_n,d'_n}\left(\psi^{\left(\frac{r_{L,n}+r_{R,n}}{2}\right)}(d_n), \psi^{\left(\frac{r_{L,n}+r_{R,n}}{2}\right)}(d'_n)\right) \\ &> \frac{1}{2}G_{d_n,d'_n}(\psi^{(r_{L,n})}(d_n),\psi^{(r_{L,n})}(d'_n)) + \frac{1}{2}G_{d_n,d'_n}(\psi^{(r_{R,n})}(d_n),\psi^{(r_{R,n})}(d'_n)) = \Gamma^*, \end{aligned}$$

since $G_{d_n,d'_n}(\cdot,\cdot)$ is strictly concave, leading to a contradiction. Therefore,

$$\psi^{(r^{L,n})}(d_n) = \psi^{(r^{R,n})}(d_n), \quad \psi^{(r^{L,n})}(d'_n) = \psi^{(r^{R,n})}(d'_n).$$

Moreover, we have $\psi^{(0)}(d_n) = \psi^{(1)}(d_n)$ and $\psi^{(0)}(d'_n) = \psi^{(1)}(d'_n)$, which completes the induction.

To conclude, we end up with a partition (P_{n_0}, Q_{n_0}) of $\mathcal{P}\times\mathcal{U}$ together with an interval $[r_{L,n_0}, r_{R,n_0}]$ for some $n_0\geq 1$. Based on the stopping criteria of the induction step, we have $\psi^{(0)}(d) = \psi^{(1)}(d)$ and $\psi^{(0)}(d') = \psi^{(1)}(d')$ for all $(d,d')\in B_{n_0}$. Furthermore, by construction, this equality is also valid for all $(d,d')\in A_{n_0}$, and consequently for all $(d,d')\in\mathcal{P}\times\mathcal{U}$, which completes the proof. □

F.2. Necessary conditions for Gaussian sampling distributions. For the specific case of Gaussian designs with known heterogeneous variance, we can further characterize the optimal allocation. In this case, the rate function equals

$$G_{d,d'}(\psi(d),\psi(d')) = \frac{(\mu^*_d - \mu^*_{d'})^2}{(\sigma^*_d)^2/\psi(d) + (\sigma^*_{d'})^2/\psi(d')}.$$

Recall that Proposition 4 yields a necessary condition that $\exists z\in\mathbb{R}^+$, for all $d\in\mathcal{P}, d'\in\mathcal{U}$,

$$\min_{\tilde{d}\in\mathcal{P}} \frac{(\mu^*_{\tilde{d}} - \mu^*_{d'})^2}{(\sigma^*_{\tilde{d}})^2/\psi(\tilde{d}) + (\sigma^*_{d'})^2/\psi(d')} = \min_{\tilde{d}'\in\mathcal{U}} \frac{(\mu^*_d - \mu^*_{\tilde{d}'})^2}{(\sigma^*_d)^2/\psi(d) + (\sigma^*_{\tilde{d}'})^2/\psi(\tilde{d}')} = z. \tag{EC.65}$$

Typically, there exists a continuum of solutions to these equations, especially when $|\mathcal{P}_c| \gg 1$ and $|\mathcal{U}_c| \gg 1$, since there are significantly more variables than equations. The solutions to the balance condition can be categorized into a combinatorial number of situations depending on whether the bivariate rate functions attain the minimum. Let $\vartheta(d,d') \in \{0,1\}$ be a binary variable, and let $\boldsymbol{\vartheta}$ be the vector of $\vartheta(d,d')$, where $d \in \mathcal{P}, d' \in \mathcal{U}$. We say that a solution to the necessary condition is of type $\boldsymbol{\vartheta}$ when $G_{d,d'}(\psi(d),\psi(d')) = \min_{\tilde{d}\in\mathcal{P}} \min_{\tilde{d}'\in\mathcal{U}} G_{\tilde{d},\tilde{d}'}(\psi(\tilde{d}),\psi(\tilde{d}'))$ if and only if $\vartheta(d,d') = 1$.

To fully characterize the optimality condition of (EC.61), we define an equivalence relation $\overset{\boldsymbol{\vartheta}}{\sim}$ on $\mathcal{D}$ for each type $\boldsymbol{\vartheta}$.

Definition EC.1. For $\tilde{d}, \bar{d} \in \mathcal{D}$, we call $\tilde{d} \overset{\boldsymbol{\vartheta}}{\sim} \bar{d}$ if there exists a chain of design pairs $\{(d_i, d_i') \in \mathcal{P} \times \mathcal{U} : i = 1, \ldots, n\}$ such that

1. $\vartheta(d_i, d_i') = 1$ for $i = 1, \ldots, n$,
2. either $d_i = d_{i+1}$ or $d_i' = d_{i+1}'$ for $i = 1, \ldots, n-1$,
3. $\tilde{d} \in \{d_1, d_1'\}$ and $\bar{d} \in \{d_n, d_n'\}$.

This equivalence relationship defines L equivalence classes of designs, i.e., $\mathcal{D} = \mathcal{D}^1 \bigcup \cdots \bigcup \mathcal{D}^L$, such that two designs $\tilde{d}, \bar{d}$ belong to the same equivalent class $\mathcal{D}^l$ if and only if $\tilde{d} \overset{\boldsymbol{\vartheta}}{\sim} \bar{d}$. Denote $\mathcal{P}^l = \mathcal{D}^l \bigcap \mathcal{P}$ and $\mathcal{U}^l = \mathcal{D}^l \bigcap \mathcal{U}$, $l = 1, \ldots, L$. Then a necessary optimality condition is as follows.

Proposition EC.2 **(Modification of Theorem 3, Zhang et al. [61])**. *The feasible solution $\psi \geq 0$ to (EC.61) is optimal only if there exist $z \in \mathcal{R}^+$ and type $\boldsymbol{\vartheta}$ with $\max_{\tilde{d}\in\mathcal{P}} \vartheta(\tilde{d}, d') = \max_{\tilde{d}'\in\mathcal{U}} \vartheta(d, \tilde{d}') = 1$, such that for each equivalent class $\mathcal{D}^l$, $l = 1, \ldots, L$, induced by $\overset{\boldsymbol{\vartheta}}{\sim}$, the following conditions hold:*

$$\sum_{d\in\mathcal{P}^l} \left(\frac{\psi(d)}{\sigma_d^*}\right)^2 = \sum_{d'\in\mathcal{U}^l} \left(\frac{\psi(d')}{\sigma_{d'}^*}\right)^2,$$

and

$$\frac{(\mu_d^* - \mu_{d'}^*)^2}{(\sigma_d^*)^2/\psi(d) + (\sigma_{d'}^*)^2/\psi(d')} \geq z,$$

with equality holds if and only if $\vartheta(d,d') = 1$.

Proof. The inequality and its equality condition follows from Proposition 4. On the other hand, recall that the complementary slackness condition (EC.29) in the KKT conditions yields

$$\Lambda_2(d,d') \left(\frac{(\mu_d^* - \mu_{d'}^*)^2}{(\sigma_d^*)^2/\psi(d) + (\sigma_{d'}^*)^2/\psi(d')} - z \right) = 0, \quad \forall d \in \mathcal{P}, d' \in \mathcal{U}.$$

Hence if $\vartheta(d,d') = 0$, then $\Lambda_2(d,d') = 0$. Since for $\ell \neq \ell'$ and $d \in \mathcal{P}^\ell$, $d' \in \mathcal{U}^{\ell'}$, the singleton $\{(d,d')\}$ per se satisfies conditions 2 and 3 for the chain in Definition EC.1, we see that $\vartheta(d,d') = 0$. Otherwise $d \overset{\boldsymbol{\vartheta}}{\sim} d'$, leading to a contradiction. This in turn implies that $\Lambda(d,d') = 0$.

Following the aforementioned facts, the first order conditions (EC.27) and (EC.28) can be rewritten as, for $l = 1, \ldots, L$,

$$\sum_{d'\in\mathcal{U}^l} \Lambda_2(d,d') \frac{\partial}{\partial x_1} G_{d,d'}(\psi(d),\psi(d')) - \Lambda_4 = 0, \quad \forall d \in \mathcal{P}^l, \tag{EC.66}$$

$$\sum_{d\in\mathcal{P}^l} \Lambda_2(d,d') \frac{\partial}{\partial x_2} G_{d,d'}(\psi(d),\psi(d')) - \Lambda_4 = 0, \quad \forall d' \in \mathcal{U}^l, \tag{EC.67}$$

where $\partial/\partial x_i$ denotes the derivative with regard to the i-th argument. It follows from the same argument in Zhang et al. [61] that the necessary conditions hold. $\square$

Remark EC.2. The conditions stated in Proposition EC.2 do not necessarily guarantee a unique solution, and thus they may not constitute a sufficient condition for optimality. The issue arises from the fact that the specific type ϑ of the optimal allocation is unknown beforehand. However, based on our understanding, what can be inferred is that the necessary conditions allow for at most one solution for each type $\boldsymbol{\vartheta}$. Note that if two solutions, denoted as $\tilde{\psi}$ and $\bar{\psi}$, are of the same type $\boldsymbol{\vartheta}$, then the inequality $\tilde{\psi}(d_0) < \bar{\psi}(d_0)$ for a certain $d_0 \in \mathcal{P}^l$, along with the balance condition and the strict monotonicity of rate functions, yields

$$\tilde{\psi}(d) < \bar{\psi}(d), \quad \text{and} \quad \tilde{\psi}(d') > \bar{\psi}(d'), \quad \forall d \in \mathcal{P}^l, d' \in \mathcal{U}^l,$$

leading to

$$\sum_{d\in\mathcal{P}^l}\left(\frac{\bar{\psi}(d)}{\sigma_d^*}\right)^2 > \sum_{d\in\mathcal{P}^l}\left(\frac{\tilde{\psi}(d)}{\sigma_d^*}\right)^2 = \sum_{d'\in\mathcal{U}^l}\left(\frac{\tilde{\psi}(d')}{\sigma_{d'}^*}\right)^2 > \sum_{d'\in\mathcal{U}^l}\left(\frac{\bar{\psi}(d')}{\sigma_{d'}^*}\right)^2,$$

a contradiction to Proposition EC.2. We conclude that the necessary optimality conditions offer at most a combinatorial number of candidates for the optimal allocation.

These necessary conditions do not trivially extend to any sampling distributions. The derivation for the optimal conditions, actually, involves the particular structure of rate functions of Gaussian designs with known variance, i.e.,

$$\frac{\partial}{\partial x_1}G_{d,d'}(\psi(d),\psi(d'))\Big/\frac{\partial}{\partial x_2}G_{d,d'}(\psi(d),\psi(d')) = f_d(\psi(d))/f_{d'}(\psi(d')),$$

where $f_d(x) = x^2/(\sigma_d^*)^2$. This "separate" property typically does not hold for the rate functions of general distribution families, such as Gaussian distributions with unknown variance, Bernoulli distributions, and Poisson distributions. For example, consider the specific case of Gaussian designs with unknown variance. In this case, the KL divergence between two Gaussian random variables with parameters $\theta_1 = (\mu_1, \sigma_1^2)$ and $\theta_2 = (\mu_2, \sigma_2^2)$ is given by

$$D(\theta_1\|\theta_2) = \ln\left(\frac{\sigma_2}{\sigma_1}\right) + \frac{\sigma_1^2 + (\mu_1-\mu_2)^2}{2\sigma_2^2} - \frac{1}{2},$$

where μ_d is the mean parameter and $\eta_d = \sigma_d^2$ is the unknown variance parameter. Hence, the rate function can be expressed as

$$\begin{aligned}
G_{d,d'}(\psi(d),\psi(d')) &= \inf_{\mu_d\le\mu_{d'}}\inf_{\sigma_d^2,\sigma_{d'}^2}\psi(d)\left(\ln\left(\frac{\sigma_d}{\sigma_d^*}\right) + \frac{(\sigma_d^*)^2 + (\mu_d^* - \mu_d)^2}{2\sigma_d^2} - \frac{1}{2}\right)\\
&\qquad + \psi(d')\left(\ln\left(\frac{\sigma_{d'}}{\sigma_{d'}^*}\right) + \frac{(\sigma_{d'}^*)^2 + (\mu_{d'}^* - \mu_{d'})^2}{2\sigma_{d'}^2} - \frac{1}{2}\right)\\
&= \inf_{\mu_d\le\mu_{d'}}\frac{1}{2}\psi(d)\ln\left(\frac{(\sigma_d^*)^2 + (\mu_d^*-\mu_d)^2}{(\sigma_d^*)^2}\right) + \frac{1}{2}\psi(d')\ln\left(\frac{(\sigma_{d'}^*)^2 + (\mu_{d'}^*-\mu_{d'})^2}{(\sigma_{d'}^*)^2}\right)\\
&= \inf_{\mu\in\mathbb{R}}\frac{1}{2}\psi(d)\ln\left(\frac{(\sigma_d^*)^2 + (\mu_d^*-\mu)^2}{(\sigma_d^*)^2}\right) + \frac{1}{2}\psi(d')\ln\left(\frac{(\sigma_{d'}^*)^2 + (\mu_{d'}^*-\mu)^2}{(\sigma_{d'}^*)^2}\right).
\end{aligned}$$

Let $\mu_{d,d'}^0$ denote a point that takes the infimum above, then by the envelop theorem [see 33, Corollary 3],

$$\begin{aligned}
\frac{\partial}{\partial x_1}G_{d,d'}(\psi(d),\psi(d')) &= \frac{1}{2}\ln\left(\frac{(\sigma_d^*)^2 + (\mu_d^* - \mu_{d,d'}^0)^2}{(\sigma_d^*)^2}\right),\\
\frac{\partial}{\partial x_2}G_{d,d'}(\psi(d),\psi(d')) &= \frac{1}{2}\ln\left(\frac{(\sigma_{d'}^*)^2 + (\mu_{d'}^* - \mu_{d,d'}^0)^2}{(\sigma_{d'}^*)^2}\right).
\end{aligned}$$

This fails to conform with the "separate" form, thus rather complicating the necessary optimality conditions and making these conditions mathematically intractable.

TABLE EC.1. The mean computation time (in seconds) when PCS first surpasses 0.9 for an increasing number of contexts.

Contexts	5	10	20	40
TTTS-C-coin	0.50	0.83	2.23	2.86

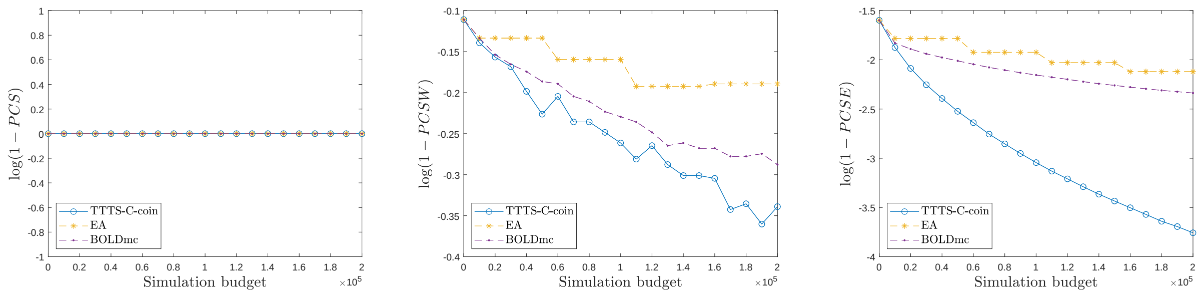

FIGURE EC.1. Log probability of incorrect selection for selecting the best design for 1,000 contexts among 50 designs per context.

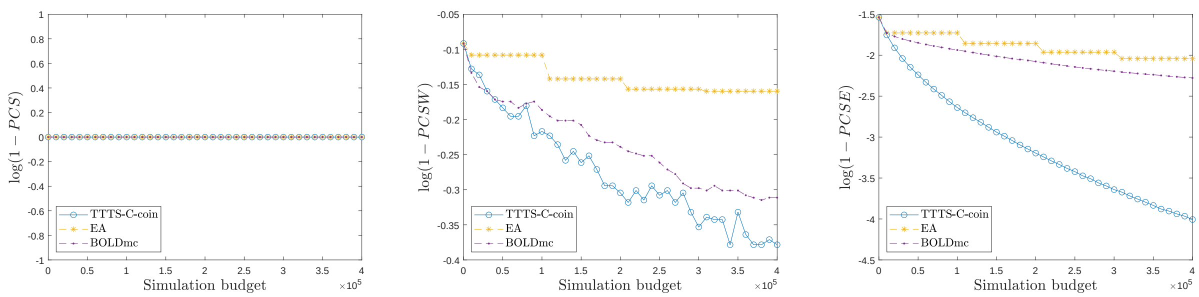

FIGURE EC.2. Log probability of incorrect selection for selecting the best design for 1,000 contexts among 100 designs per context.

Appendix G: Supplementary experiments

G.1. Computation time Following the Gaussian example in Section 6.1, we conduct 4 synthetic experiments with $5, 10, 20, 40$ contexts, respectively. In each experiment, we estimate the CPU time in need to achieve a level of PCS larger than 0.9 using 1,000 macro-replications. These experiments are implemented in Matlab on a laptop with an AMD Ryzen 7 5800H CPU. Table EC.1 illustrates that the computation time grows sub-linearly. The average computation time per context using TTTS-C-coin, which utilizes a fixed hyperparameter γ, is inclined to decrease with respect to the number of contexts. This is consistent with the claim that the extra computation time as the number of contexts increases is marginal.

G.2. Scaling the number of contexts Following Section 6.1, we supplement experiments under a Gaussian setting. We consider two synthetic examples with $|\mathcal{C}| = 1,000$. The numbers of designs in each context are 50 and 100, respectively, in these two examples. Again the mean and standard deviation of each design are generated from $N(0, 10)$ and $\text{Unif}([4, 6])$. The compared algorithms are TTTS-C-coin, EA and BOLDmc. AOAmc is excluded in comparison due to the prohibitively long computation time of implementing macro-replications.

In both examples, the performance of TTTS-C-coin is significantly superior to EA and BOLDmc in terms of PCSE. When the simulation budget is exhausted, TTTS-C-coin yields a PCSE over 97.5%, whereas the PCSE of EA and BOLDmc are below 92%. In terms of the metric of PCSW, TTTS-C-coin demonstrates at least an improvement of 4% over EA and BOLDmc. Since the simulation budget of 200,000 samples and 400,000 samples in these two examples, respectively, equates to an average of 4 samples allocated to each design, the budget is insufficient to yield a positive probability of simultaneous correct selection. As a result, all algorithms lead to zero PCS.