Online Appendix
Design-Based Confidence Sequences: A General Approach to Risk Mitigation in Panel Experiments

Dae Woong Ham, Michael Lindon, Martin Tingley, and Iavor Bojinov

A Example for multi-arm bandit

We first remark that our propensity scores in Lemma 4.2 for the n units, $p_{i|i-1}(w)$, can be fully adaptive. This now allows us to tackle the common bandit settings, where an agent is typically tasked to maximize reward and pull the next arm as a function of all the previous treatment assignments (Slivkins, 2019; Sutton and Barto, 2018). Although the positivity assumption in Assumption 2.1 for time series and panel experiments is more likely to hold and controlled by the experiment, in adaptive bandit settings Assumption 2.1 may not hold asymptotically if the adaptive algorithm converges to pulling an optimal arm with probability one. Therefore, this assumption is often a "stricter" assumption for multi-arm bandits as many adaptive algorithms such as Thompson Sampling (Thompson, 1933; Russo, 2020) make the adaptive probabilities converge to one or zero. To solve this concern, we recommend practitioners to use a mixture of a uniform explore policy and an adaptive exploit policy, e.g., assigning an epsilon mixture weight to the uniform explore policy ensures the propensity scores are bounded. We now show an example of using Lemma 4.2 for a simplified two arm bandit setting.

Example A.1 (Two Arm Bandit)**.** Suppose for simplicity the two-arm bandit problem, where an agent pulls either arm A (control) or arm B (treatment). Suppose the rewards under arms A and B have distributions $N(1, 1)$ and $N(2, 1)$, respectively. Consider the following adaptive probabilities for the n^{th} individual

$$p_{n|n-1}(1) = \frac{\bar{Y}_{1,n-1}}{\bar{Y}_{1,n-1} + \bar{Y}_{0,n-1}}, \quad n > 10$$

and $p_{n|n-1}(0) = 1 - p_{n|n-1}(1)$, where $\bar{Y}_{1,n-1}$ is the sample mean of arm B using the realized samples for $i = 1, 2, \ldots, n-1$ and $\bar{Y}_{0,n-1}$ is defined similarly. In other words, the agent upweights the arm that produce a higher mean reward based on the sample means. Lastly, the agent flips a fair coin for the first ten time periods (exploration period). Figure 7 shows that even under adaptively sampled data, our confidence sequence tightens to the desired truth. For this case, the agent would likely terminate the experiment at approximately $n = 150$ because the reward from arm B is statistically significantly higher than the reward from arm A.

B Proof of Theorem 2.1

$$E(\hat{\tau}_t \mid \mathcal{F}_{t-1}) = \tau_t(w^{obs}_{1:(t-1)})$$

Proof.

$$\begin{aligned} E(\hat{\tau}_t \mid \mathcal{F}_{t-1}) &= E\left(\frac{\mathbb{1}\{W_t = 1\}Y_t}{p_{t|t-1}(1)} - \frac{\mathbb{1}\{W_t = 0\}Y_t}{p_{t|t-1}(0)} \mid \mathcal{F}_{t-1}\right) \\ &= \frac{p_{t|t-1}(1)Y_t(w^{obs}_{1:(t-1)}, 1)}{p_{t|t-1}(1)} - \frac{p_{t|t-1}(0)Y_t(w^{obs}_{1:(t-1)}, 0)}{p_{t|t-1}(0)} \\ &= \tau_t(w^{obs}_{1:(t-1)}). \end{aligned}$$

□

Next, we calculate the closed form expression of $\mathrm{Var}(\hat{\tau}_{t|t-1} \mid \mathcal{F}_{t-1})$.

$$
\begin{aligned}
\mathrm{Var}(\hat{\tau}_{t|t-1} \mid \mathcal{F}_{t-1}) &= \mathrm{Var}\left(\frac{\mathbb{1}\{W_t = 1\}Y_t}{p_{t|t-1}(1)} - \frac{\mathbb{1}\{W_t = 0\}Y_t}{p_{t|t-1}(0)} \mid \mathcal{F}_{t-1}\right) \\
&= \frac{p_{t|t-1}(1)p_{t|t-1}(0)Y_t(w^{obs}_{1:(t-1)}, 1)^2}{p_{t|t-1}(1)^2} + \frac{p_{t|t-1}(0)p_{t|t-1}(1)Y_t(w^{obs}_{1:(t-1)}, 0)^2}{p_{t|t-1}(0)^2} \\
&+ 2Y_t(w^{obs}_{1:(t-1)}, 1)Y_t(w^{obs}_{1:(t-1)}, 0) \\
&= \frac{p_{t|t-1}(0)Y_t(w^{obs}_{1:(t-1)}, 1)^2}{p_{t|t-1}(1)} + \frac{p_{t|t-1}(1)Y_t(w^{obs}_{1:(t-1)}, 0)^2}{p_{t|t-1}(0)} + 2Y_t(w^{obs}_{1:(t-1)}, 1)Y_t(w^{obs}_{1:(t-1)}, 0) \\
&= \frac{\Big(p_{t|t-1}(0)Y_t(w^{obs}_{1:(t-1)}, 1) + p_{t|t-1}(1)Y_t(w^{obs}_{1:(t-1)}, 0)\Big)^2}{p_{t|t-1}(1)p_{t|t-1}(0)},
\end{aligned}
$$

where the third line follows because $\mathrm{Cov}(\mathbb{1}\{W_t = 1\}, \mathbb{1}\{W_t = 0\} \mid \mathcal{F}_{t-1}) = -p_{t|t-1}(1)p_{t|t-1}(0)$. Now we show that

$$
\mathrm{Var}(\hat{\tau}_{t|t-1} \mid \mathcal{F}_{t-1}) \le \sigma_{t|t-1} := \frac{Y_t(w^{obs}_{1:(t-1)}, 1)^2}{p_{t|t-1}(1)} + \frac{Y_t(w^{obs}_{1:(t-1)}, 0)^2}{p_{t|t-1}(0)},
$$

which would complete the proof because it is straight forward to show that $E(\hat{\sigma}_{t|t-1} \mid \mathcal{F}_{t-1}) = \sigma_{t|t-1}$.

Proof.

$$
\begin{aligned}
\mathrm{Var}(\hat{\tau}_{t|t-1} \mid \mathcal{F}_{t-1}) &= \frac{p_{t|t-1}(0)^2Y_t(w^{obs}_{1:(t-1)}, 1)^2 + p_{t|t-1}(1)^2Y_t(w^{obs}_{1:(t-1)}, 0)^2}{p_{t|t-1}(1)p_{t|t-1}(0)} \\
&+ \frac{2p_{t|t-1}(0)Y_t(w^{obs}_{1:(t-1)}, 1)p_{t|t-1}(1)Y_t(w^{obs}_{1:(t-1)}, 0)}{p_{t|t-1}(1)p_{t|t-1}(0)} \\
&\le \frac{Y_t(w^{obs}_{1:(t-1)}, 1)^2}{p_{t|t-1}(1)} + \frac{Y_t(w^{obs}_{1:(t-1)}, 0)^2}{p_{t|t-1}(0)},
\end{aligned}
$$

where the last line follows because $(a-b)^2 \ge 0 \to a^2 + b^2 \ge 2ab$ and we let $a = p_{t|t-1}(0)Y_t(w^{obs}_{1:(t-1)}, 1)$ and $b = p_{t|t-1}(1)Y_t(w^{obs}_{1:(t-1)}, 0)$. □

Lastly, $\gamma^2_{t|t-1}$ defined in Assumption 5.1 can be obtained by replacing $Y_t(w^{obs}_{1:(t-1)}, \bullet)$ with a new "residualized" potential outcome $\tilde{Y}_t(w^{obs}_{1:(t-1)}, \bullet) := Y_t(w^{obs}_{1:(t-1)}, \bullet) - \hat{Y}_{t|t-1}$. Since all the proof is conditioned on the filtration, the above is equivalent to introducing a constant and therefore we have that

$$
\gamma^2_{t|t-1} = \frac{\tilde{Y}_t(w^{obs}_{1:(t-1)}, 1)^2}{p_{t|t-1}(1)} + \frac{\tilde{Y}_t(w^{obs}_{1:(t-1)}, 0)^2}{p_{t|t-1}(0)}. \tag{12}
$$

C Alternative closed-form exact design-based confidence sequence

To derive an alternative exact confidence sequence, we apply the empirical Bernstein inequalities to derive a design-based confidence sequence (Maurer and Pontil, 2009; Fan, Grama and Liu, 2015a).

Theorem C.1 (Design-based Non-asymptotic Confidence Sequence for the CTE)**.** Suppose $\{W_t, Y_t\}_{t=1}^T$ are observed for arbitrary data dependent stopping time T, where Assumptions 2.1 and 2.2 are satisfied. Then,

$$\frac{1}{t}\sum_{j=1}^{t}\hat{\tau}_{j|j-1} \pm \left[\frac{m(m+1)}{t}\log\left(\frac{2}{\alpha}\right) + \frac{S_{t|t-1}}{t}\left(\frac{m+1}{m}\log\left(1+\frac{1}{m}\right) - \frac{1}{m}\right)\right]$$

forms a valid $(1-\alpha)$ confidence sequence for the running mean of the contemporaneous average treatment effect $\bar{\tau}_t(w^{obs}_{1:(t-1)})$.

The proof is provided in the bottom of this Appendix section. The key part of the proof is that we show

$$\exp\left[\frac{\sum_{j=1}^{t}\hat{\tau}_{j|j-1} - \tau_t(w^{obs}_{1:(t-1)})}{m(m+1)} + \frac{S_{t|t-1}}{m^2}\left(\log\left(\frac{m}{m+1}\right) + \frac{1}{m+1}\right)\right]$$

forms a supermartingale by applying Fan's inequality (Fan, Grama and Liu, 2015b) through similar proof techniques in (Howard et al., 2020). The rest follows from algebraically manipulating the supermartingale after applying Lemma 3.1.

We note that the confidence width is of order approximately $S_{t|t-1}/t$, which does not shrink asymptotically ($t \to \infty$) to zero unless $S_{t|t-1}$ grows sub-linearly ($\hat{\sigma}_t$ vanishes to zero) or there are stronger assumptions on the potential outcomes. This issue was recognized in Howard et al. (2020) who propose a solution under additional conditions. Therefore, we similarly propose Theorem 3.2 instead. We still present a simulation figure using this exact CS Example C.1 below.

Before stating the proof we quickly show an example and simulation to demonstrate this confidence sequence.

Example C.1 (Simple Messaging Alerting Treatment)**.** Suppose the same scenario as Example 4.1 but simplified to a stationary constant treatment effect, i.e., no novelty effect but there is a constant treatment effect of 0.10 across time. Then, in the notation from Theorem C.1, $M = 1$ and $m = 1/0.5 = 2$. Furthermore, the the running mean of the CTE is roughly -0.1 at any time t. Although this scenario is largely simplified to remove any non-stationary, i.e., the user responds exactly the same way to a message alert at any time t, we present it this way to pedagogically demonstrate the theoretical properties of Theorem C.1. The left and right panels of Figure 8 show the confidence sequence from applying Theorem C.1 on one simulated experimental data in the aforementioned setting for $T = 500$ and $T = 10,000$ time points, respectively. The left panel shows that the analyst would likely stop at the $t = 210^{\text{th}}$ time because the confidence sequence shows a statistically significant negative treatment effect (the first time the red contours do not overlap the black dotted line), mitigating further harm. However, the right panel shows that the confidence width does not shrink despite a large $T = 10,000$, illustrating the aforementioned limitations.

We now prove Theorem C.1. We build off the proof of Theorem 4 in (Howard et al., 2020) and apply it to our specific time series setting. As hinted in Section 3.2, we will first show that

$$\exp\left[\frac{\sum_{j=1}^{t}\hat{\tau}_{j|j-1} - \tau_t(w^{obs}_{1:(t-1)})}{m(m+1)} + \frac{S_{t|t-1}}{m^2}\left(\log\left(\frac{m}{m+1}\right) + \frac{1}{m+1}\right)\right]$$

is a non-negative supermartingale with respect to the filtration $\mathcal{F}_{N,n-1}$. (Fan, Grama and Liu, 2015b) show that

$$\exp\left(\lambda\kappa + \kappa^2(\lambda + \log(1-\lambda))\right) \leq 1 + \lambda\kappa$$

for $\kappa \geq -1$ and $\lambda \in [0, 1)$ from Fan's Inequality. We let

$$\kappa = \frac{\hat{\tau}_{t|t-1}}{m},$$

where $\kappa \geq -1$ since $|\hat{\tau}_{t|t-1}| \leq m$ for every t by Assumption 2.2. Therefore, we have

$$\exp\left(\lambda\frac{\hat{\tau}_{t|t-1}}{m} + \frac{\hat{\tau}^2_{t|t-1}}{m^2}(\lambda + \log(1-\lambda))\right) \leq 1 + \lambda\frac{\hat{\tau}_{t|t-1}}{m}$$

$$E\left[\exp\left(\frac{\lambda\hat{\tau}_{t|t-1}}{m} + \frac{\hat{\sigma}^2_{t|t-1}}{m^2}(\lambda + \log(1-\lambda))\right) \mid \mathcal{F}_{t-1}\right] \leq 1 + \lambda\frac{\tau_t(w^{obs}_{1:(t-1)})}{m}$$

$$E\left[\exp\left(\frac{\lambda(\hat{\tau}_{t|t-1} - \tau_t(w^{obs}_{1:(t-1)}))}{m} + \frac{\hat{\sigma}^2_{t|t-1}}{m^2}(\lambda + \log(1-\lambda))\right) \mid \mathcal{F}_{t-1}\right] \leq \exp(-\frac{\lambda\tau_t(w^{obs}_{1:(t-1)})}{m})[1 + \lambda\frac{\tau_t(w^{obs}_{1:(t-1)})}{m}]$$

$$E\left[\exp\left(\frac{\lambda(\hat{\tau}_{t|t-1} - \tau_t(w^{obs}_{1:(t-1)}))}{m} + \frac{\hat{\sigma}^2_{t|t-1}}{m^2}(\lambda + \log(1-\lambda))\right) \mid \mathcal{F}_{t-1}\right] \leq 1,$$

where the second line follows because $\hat{\tau}^2_{t|t-1} = \hat{\sigma}^2_{t|t-1}$ and Theorem 2.1 and the last line follows because $1 + x \leq \exp(x)$. We plug $\lambda = 1/(m+1)$ and because the above is a non-negative quantity. This directly implies that

$$\exp\left[\frac{\sum_{j=1}^t \hat{\tau}_{j|j-1} - \tau_t(w^{obs}_{1:(t-1)})}{m(m+1)} + \frac{S_{t|t-1}}{m^2}\left(\log\left(\frac{m}{m+1}\right) + \frac{1}{m+1}\right)\right]$$

is indeed a non-negative super martingale with respect to $\mathcal{F}_{N,n-1}$ as desired with initial value less than one. We apply Lemma 3.1 and have that

$$\Pr\left(\exists t\,:\, \exp\left[\frac{\sum_{j=1}^t(\hat{\tau}_{j|j-1} - \tau_t(w^{obs}_{1:(t-1)}))}{m(m+1)} + \frac{S_{t|t-1}}{m^2}\left(\log\left(\frac{m}{m+1}\right) + \frac{1}{m+1}\right)\right] \geq \frac{1}{\tilde{\alpha}}\right) \leq \tilde{\alpha}$$

$$\Pr\left(\exists t\,:\, \left[\frac{\sum_{j=1}^t(\hat{\tau}_{j|j-1} - \tau_t(w^{obs}_{1:(t-1)}))}{m(m+1)} + \frac{S_{t|t-1}}{m^2}\left(\log\left(\frac{m}{m+1}\right) + \frac{1}{m+1}\right)\right] \geq \log\left(\frac{1}{\tilde{\alpha}}\right)\right) \leq \tilde{\alpha}$$

$$\Pr\left(\exists t\,:\, \sum_{j=1}^t(\hat{\tau}_{j|j-1} - \tau_t(w^{obs}_{1:(t-1)})) \geq m(m+1)\log\left(\frac{1}{\tilde{\alpha}}\right) - \frac{(m+1)S_{t|t-1}}{m}\left(\log\left(\frac{m}{m+1}\right) + \frac{1}{m+1}\right)\right) \leq \tilde{\alpha}$$

$$\Pr\left(\exists t\,:\, \sum_{j=1}^t(\hat{\tau}_{j|j-1} - \tau_t(w^{obs}_{1:(t-1)})) \geq \left[m(m+1)\log\left(\frac{1}{\tilde{\alpha}}\right) + S_{t|t-1}\left(\frac{m+1}{m}\log\left(1+\frac{1}{m}\right) - \frac{1}{m}\right)\right]\right) \leq \tilde{\alpha}.$$

Consequently, we have that

$$\Pr\left(\exists t\,:\, \frac{1}{t}\sum_{j=1}^t(\hat{\tau}_{j|j-1} - \tau_t(w^{obs}_{1:(t-1)})) \geq \left[\frac{m(m+1)}{t}\log\left(\frac{1}{\tilde{\alpha}}\right) + \frac{S_{t|t-1}}{t}\left(\frac{m+1}{m}\log\left(1+\frac{1}{m}\right) - \frac{1}{m}\right)\right]\right) \leq \tilde{\alpha}.$$

This gives the one-sided confidence sequence and we can do the same trick and build the same statement instead for $\kappa = -\hat{\tau}_{t|t-1}/m$. Then we get

$$\Pr\left(\exists t\,:\, -\frac{1}{t}\sum_{j=1}^t(\hat{\tau}_{j|j-1} - \tau_t(w^{obs}_{1:(t-1)})) \geq \left[\frac{m(m+1)}{t}\log\left(\frac{1}{\tilde{\alpha}}\right) + \frac{S_{t|t-1}}{t}\left(\frac{m+1}{m}\log\left(1+\frac{1}{m}\right) - \frac{1}{m}\right)\right]\right) \leq \tilde{\alpha}.$$

Taking $\alpha = \tilde{\alpha}/2$ and applying the union bound completes the proof.

D Proof of Theorem 3.2

The above proof shows that

$$M_t := \exp\left(\lambda A_t + B_t(\lambda + \log(1-\lambda))\right) \tag{13}$$

is a super-martingale with initial value 1, where

$$A_t := \frac{\sum_{j=1}^{t} \hat{\tau}_{t|t-1} - \tau_t(w^{obs}_{1:(t-1)})}{m}, \quad B_t := \frac{S_{t|t-1}}{m^2}.$$

For any distribution F on $(0, 1)$, we have by Fubini's theorem that

$$\tilde{M}_t := \int_{\lambda \in (0,1)} M_t dF(\lambda)$$

is again another super-martingale with initial value 1. Following the proof of Theorem 2 in (Waudby-Smith et al., 2022), we choose the truncated gamma distribution given by

$$f(\lambda) = \frac{\delta^\delta e^{-\delta(1-\lambda)} (1-\lambda)^{\delta-1}}{\Gamma(\delta) - \Gamma(\delta, \delta)}$$

for any $\delta \geq 0$. Therefore, we have that

$$\begin{aligned}
\tilde{M}_t &= \int_0^1 \exp\left\{\lambda A_t + B_t(\lambda + \log(1-\lambda)\right\} f(\lambda) d\lambda \\
&= \int_0^1 \exp\left\{\lambda A_t + B_t(\lambda + \log(1-\lambda)\right\} \frac{\delta^\delta e^{-\delta(1-\lambda)} (1-\lambda)^{\delta-1}}{\Gamma(\delta) - \Gamma(\delta, \delta)} d\lambda \\
&= \frac{\delta^\delta e^{-\delta}}{\Gamma(\delta) - \Gamma(\delta, \delta)} \int_0^1 \exp\{\lambda \left(A_t + B_t + \delta\right)\} (1-\lambda)^{B_t + \delta - 1} d\lambda \\
&= \left(\frac{\delta^\delta e^{-\delta}}{\Gamma(\delta) - \Gamma(\delta, \delta)}\right) \left(\frac{1}{B_t + \delta}\right) {}_1F_1(1, B_t + \delta + 1, A_t + B_t + \delta),
\end{aligned}$$

where the last line follows from the definition of the Kummer's confluent hypergeometric function.

Therefore, we have by Lemma 3.1 that

$$\Pr\left(\exists t : \left(\frac{\delta^\delta e^{-\delta}}{\Gamma(\delta) - \Gamma(\delta, \delta)}\right) \left(\frac{1}{B_t + \delta}\right) {}_1F_1(1, B_t + \delta + 1, A_t + B_t + \delta) \geq \frac{1}{\tilde{\alpha}}\right) \leq \tilde{\alpha}.$$

Consequently, a one-sided lower confidence sequence, i.e., a confidence sequence that is bounded from below but covers up to infinity, can be obtained by a root-finding algorithm to find all

$$\{\tau_t(w^{obs}_{1:(t-1)}) : V_t(\tau_t(w^{obs}_{1:(t-1)}) \geq \frac{1}{\tilde{\alpha}}\},$$

where

$$V_t(\tau_t(w^{obs}_{1:(t-1)})) := \left(\frac{\delta^\delta e^{-\delta}}{\Gamma(\delta) - \Gamma(\delta, \delta)}\right) \left(\frac{1}{B_t + \delta}\right) {}_1F_1(1, B_t + \delta + 1, A_t + B_t + \delta).$$

An upper confidence sequence can be obtained in a similar way by using the mirror trick and replacing $\tau_t(w^{obs}_{1:(t-1)})$ with $1 - \tau_t(w^{obs}_{1:(t-1)})$. Furthermore, this confidence sequence does not solve the issue where it requires the analyst to know M and p_{min} before the experiment. Additionally, this confidence sequence does not have a closed-form expression, thus it requires a root-solving algorithm to build the confidence sequence. However, (Waudby-Smith et al., 2022) show that this provably has an asymptotic rate of $O(\sqrt{B_t \log(B_t)}/t)$, which does solve the issue related to the order of the confidence sequence width.

E Proof of Theorem 4.1

The proof proceeds in three steps. We note that this proof leverages the proof of Theorem 2.3 in (Waudby-Smith et al., 2021) but extended to our setting.

Step 1: Building martingale using Gaussian distribution Recently, (Ramdas et al., 2020) shows that all sequential tests must have an explicit or implicit construction of a non-negative martingale. Although one of the major advantages of an asymptotic confidence sequences is that it avoids explicitly constructing a martingale, the proof still relies on constructing a martingale with the asymptotic Gaussian distribution. Consequently, the first step of the proof builds a martingale from a sequence of iid standard Gaussian random variables.

Let $(Z_t)_{t=1}^{\infty}$ be a sequence of iid standard Gaussian random variable. We note that

$$M_t(\lambda) := \exp\left(\sum_{j=1}^{t}(\lambda\sigma_{j|j-1}Z_j - \lambda^2\sigma_{j|j-1}^2/2)\right)$$

is a non-negative martingale starting at one for any $\lambda \in \mathbb{R}$ with respect to the canonical filtration (Robbins, 1970). For algebraic simplicity, we also define $L_t := \sum_{j=1}^{t}\sigma_{j|j-1}Z_j$ and $\bar{\sigma}_t^2 = \frac{1}{t}\sum_{j=1}^{t}\sigma_{j|j-1}^2$. Moreover, for any probability distribution $F(\lambda)$ on $\mathbb{R}$, we also have the mixture,

$$\int_{\lambda\in\mathbb{R}} M_t(\lambda)dF(\lambda)$$

is again a non-negative martingale with initial value one (Robbins, 1970). In particular, we consider the probability distribution function $f(\lambda; 0, \eta^2)$ for the Gaussian distribution with mean zero and variance η^2 as the mixing distribution. The resulting martingale is

$$\begin{aligned}
M_t :=& \int_{\lambda\in\mathbb{R}} M_t(\lambda)f(\lambda;0,\eta^2)d\lambda \\
=&\frac{1}{\sqrt{2\pi\eta^2}}\int_\lambda \exp\left(\lambda L_t - \frac{t\lambda^2\bar{\sigma}_t^2}{2}\right)\exp\left(\frac{-\lambda^2}{2\eta^2}\right)d\lambda \\
=&\frac{1}{\sqrt{2\pi\eta^2}}\int_\lambda \exp\left(\lambda L_t - \frac{\lambda^2(1+t\eta^2\bar{\sigma}_t^2)}{2\eta^2}\right)d\lambda \\
=&\frac{1}{\sqrt{2\pi\eta^2}}\int_\lambda \exp\left(\frac{-\lambda^2(1+t\eta^2\bar{\sigma}_t^2)+2\lambda\eta^2 L_t}{2\eta^2}\right)d\lambda \\
=&\frac{1}{\sqrt{2\pi\eta^2}}\int_\lambda \exp\left(\frac{-a(\lambda^2+\frac{b}{a}2\lambda}{2\eta^2}\right)d\lambda,
\end{aligned}$$

where $a = t\eta^2\bar{\sigma}_t^2 + 1$ and $b = \eta^2 L_t$. Completing the square, we have that the integrand is:

$$\exp\left(\frac{-a(\lambda^2+\frac{b}{a}2\lambda}{2\eta^2}\right) = \exp\left(\frac{-(\lambda-b/a)^2}{2\eta^2/a}\right)\exp\left(\frac{b^2}{2a\eta^2}\right).$$

Putting the expression back into M_t we have that,

$$M_t = \frac{1}{\sqrt{2\pi\eta^2/a}} \int_\lambda \exp\left(\frac{-(\lambda - b/a)^2}{2\eta^2/a}\right) d\lambda \frac{\exp\left(\frac{b^2}{2a\eta^2}\right)}{\sqrt{a}}$$
$$= \exp\left(\frac{\eta^2(\sum_{j=1}^t \sigma_{j|j-1} Z_j)^2}{2(t\bar{\sigma}_t^2\eta^2 + 1)}\right)(t\bar{\sigma}_t^2\eta^2 + 1)^{-1/2},$$

where the last line follows because the first part of the first line is one and we plug back in the definition of a and b.

Since M_t is a non-negative martingale with initial value one, we can use Lemma 3.1 to claim that

$$\Pr(\forall t \geq 1, M_t < 1/\alpha) \geq 1 - \alpha$$
$$= \Pr\left(\forall t \geq 1, \left|\frac{1}{t}\sum_{j=1}^t \sigma_{j|j-1} Z_j\right| < \sqrt{\frac{2(t\bar{\sigma}_t^2\eta^2 + 1)}{t^2\eta^2} \log\left(\frac{\sqrt{t\bar{\sigma}_t^2\eta^2 + 1}}{\alpha}\right)}\right) \geq 1 - \alpha, \tag{14}$$

where the last line follows from taking the logarithm and simple algebraic manipulation.

Step 2: Strong Approximation via Martingale Sequence Differences We first define the estimation error as

$$u_t = \hat{\tau}_{t|t-1} - \tau_t(w^{obs}_{1:(t-1)}).$$

By Theorem 2.1, $\{u_t\}$ is a martingale difference sequence with respect to $\mathcal{F}_{t-1}$. Similar to the proof of Step 2 of (Waudby-Smith et al., 2021), we also use the strong approximation theorem presented in (Strassen, 1967). In particular, we require Equation (159) in Theorem 4.4 of Strassen's paper (further details in Lemma A.3 of (Waudby-Smith et al., 2021)) for our strong approximation theorem. However, our proof is different than that in Step 2 of (Waudby-Smith et al., 2021) for the following reason.

The original Theorem 4.4 in (Strassen, 1967) is stated for martingales difference sequence of the form $E(D_n \mid \sigma(D_1, \ldots, D_{n-1})) = 0$, where $D_i := (W_i, Y_i)$ is the data at time i. Although our martingale is of the form $E(f(D_n) \mid \sigma(D_1, \ldots, D_{n-1})) = 0$, where $f(.)$ is the function that maps the data to $\hat{\tau}_{t|t-1} - \tau_t(w^{obs}_{1:(t-1)})$. These slight differences can complicate the original proof or Strassen's Theorem 4.4 depending on the function $f(.)$. Regardless, we can still apply Theorem 4.4 by defining a larger filtration $\tilde{\mathcal{F}}_t := \sigma(u_1, u_2, \ldots, u_t)$ and apply Strassen's theorem directly to a MDS of the form $E(u_t \mid \tilde{\mathcal{F}}_{t-1})$. Note our estimand do not change even if conditioning on a larger filtration because we are in the design-based framework (hence these causal estimands are fixed regardless of what random quantity one conditions on). Regardless, this step is just a trick to apply Strassen's theorem faithfully. Specifically, we can still write this proof with respect to the original filtration $\sigma(D_1, \ldots, D_{n-1})$ since u_t is still a MDS with respect to this smaller filtration (once we have gotten the strong approximation we can apply the same trick as outlined below).

The remaining conditions are identical and we omit the uniform integrability condition in Equation (138) of (Strassen, 1967) since it holds trivially under our bounded potential outcome for Assumption 2.2 and our strong positivity in Assumption 2.2. Lastly, although this theorem uses the actual variance (not an upper bound), using an upper bound only makes the confidence sequence width strictly wider and hence the validity still holds.

Finally, utilizing Theorem 4.4 Equation (159) in (Strassen, 1967), we have that

$$\frac{1}{t}\sum_{j=1}^t u_j = \frac{1}{t}\sum_{j=1}^t \sigma_{j|j-1} Z_j + o\left(\frac{\tilde{S}_{t|t-1}^{3/8} \log(\tilde{S}_{t|t-1})}{t}\right) \quad a.s., \tag{15}$$

Combining Equation (14) and Equation (15) implies that with probability at least $(1-\alpha)$,

$$\forall t \geq 1, \left|\frac{1}{t}\sum_{i=1}^{t} u_i\right| < \sqrt{\frac{2(t\bar{\sigma}_t^2\eta^2+1)}{t^2\eta^2}\log\left(\frac{\sqrt{t\bar{\sigma}_t^2\eta^2+1}}{\alpha}\right)} + o\left(\frac{\tilde{S}_{t|t-1}^{3/8}\log(\tilde{S}_{t|t-1})}{t}\right). \tag{16}$$

Using Assumption 4.1, we have that

$$\frac{1}{t}\sum_{j=1}^{t}\hat{\tau}_{j|j-1} \pm \sqrt{\frac{2(t\bar{\sigma}_t^2\eta^2+1)}{t^2\eta^2}\log\left(\frac{\sqrt{t\bar{\sigma}_t^2\eta^2+1}}{\alpha}\right)} \tag{17}$$

forms an $(1-\alpha)-$asymptotic confidence sequence for $\bar{\tau}_t(w^{obs}_{1:(t-1)})$, where we used Assumption 4.1 so that the $\tilde{V}_t/V_t \xrightarrow{a.s.} 1$ holds where $\tilde{V}_t$ is the non-asymptotic confidence width in Equation (16) (with the little o term) and V_t is defined in Equation (17) (without the little o term).

Step 3: Using empirical variance Unfortunately, the confidence sequence in Equation (17) can not be directly used because $\bar{\sigma}_t$ is based off the true variance and hence not obtainable from the data. The last step is to replace Equation (17) with our estimated variance $\tilde{\sigma}_t^2 := S_{t|t-1}/t$.

More precisely, we now show that if we further have $\tilde{\sigma}_t^2 \xrightarrow{a.s.} \bar{\sigma}_t^2$, then we have

$$\frac{1}{t}\sum_{j=1}^{t}\hat{\tau}_{j|j-1} \pm \sqrt{\frac{2(t\tilde{\sigma}_t^2\eta^2+1)}{t^2\eta^2}\log\left(\frac{\sqrt{t\tilde{\sigma}_t^2\eta^2+1}}{\alpha}\right)}$$

forms a $(1-\alpha)$-asymptotic confidence sequence for $\tau_t(w^{obs}_{1:(t-1)})$, giving us the desired result. For completeness, we replicate this part of the proof under our setting. First we rewrite the assumption of $\tilde{\sigma}_t^2 \xrightarrow{a.s.} \bar{\sigma}_t^2$ as $\tilde{\sigma}_t^2 - \bar{\sigma}_t^2 = o(\bar{\sigma}_t^2)$. Then Equation (17) gives us

$$\begin{aligned}
\sqrt{\frac{2(t\bar{\sigma}_t^2\eta^2+1)}{t^2\eta^2}\log\left(\frac{\sqrt{t\bar{\sigma}_t^2\eta^2+1}}{\alpha}\right)} &= \sqrt{\frac{2(t(\tilde{\sigma}_t^2+o(\bar{\sigma}_t^2))\eta^2+1)}{t^2\eta^2}\log\left(\frac{\sqrt{t(\tilde{\sigma}_t^2+o(\bar{\sigma}_t^2))\eta^2+1}}{\alpha}\right)} \\
&= \sqrt{\frac{t(\tilde{\sigma}_t^2+o(\bar{\sigma}_t^2))\eta^2+1}{t^2\eta^2}\log\left(\frac{t(\tilde{\sigma}_t^2+o(\bar{\sigma}_t^2))\eta^2+1}{\alpha^2}\right)} \\
&= \sqrt{\frac{t\tilde{\sigma}_t^2\eta^2+o(t\bar{\sigma}_t^2)+1}{t^2\eta^2}\log\left(\frac{t\tilde{\sigma}_t^2\eta^2+o(t\bar{\sigma}_t^2)+1}{\alpha^2}\right)} \\
&= \sqrt{\left(\frac{t\tilde{\sigma}_t^2\eta^2+1}{t^2\eta^2}+o(\bar{\sigma}_t^2/t)\right)\log\left(\frac{t\tilde{\sigma}_t^2\eta^2+o(t\bar{\sigma}_t^2)+1}{\alpha^2}\right)}.
\end{aligned}$$

Focusing on the second logarithmic term, we have

$$\begin{aligned}
\log\left(\frac{t\tilde{\sigma}_t^2\eta^2 + o(t\bar{\sigma}_t^2) + 1}{\alpha^2}\right) &= \log\left(\frac{t\tilde{\sigma}_t^2\eta^2 + 1}{\alpha^2} + o(t\bar{\sigma}_t^2)\right) \\
&= \log\left(\frac{t\tilde{\sigma}_t^2\eta^2 + 1}{\alpha^2}\left[1 + o(1)\right]\right) \\
&= \log\left(\frac{t\tilde{\sigma}_t^2\eta^2 + 1}{\alpha^2}\right) + \log(1 + o(1)) \\
&= \log\left(\frac{t\tilde{\sigma}_t^2\eta^2 + 1}{\alpha^2}\right) + o(1),
\end{aligned}$$

where the last line follows because $\log(1+x) = x + o(1)$ for $|x| < 1$. Returning back to the main expression we have

$$\begin{aligned}
\sqrt{\frac{2(t\bar{\sigma}_t^2\eta^2+1)}{t^2\eta^2}\log\left(\frac{\sqrt{t\bar{\sigma}_t^2\eta^2+1}}{\alpha}\right)} &= \sqrt{\left(\frac{t\tilde{\sigma}_t^2\eta^2+1}{t^2\eta^2} + o(V_t/t^2)\right)\left[\log\left(\frac{t\tilde{\sigma}_t^2\eta^2+1}{\alpha^2}\right) + o(1)\right]} \\
&= \sqrt{\frac{t\tilde{\sigma}_t^2\eta^2+1}{t^2\eta^2}\log\left(\frac{t\tilde{\sigma}_t^2\eta^2+1}{\alpha^2}\right) + o(V_t/t^2) + o(V_t\log V_t/t^2) + o(V_t/t^2)} \\
&= \sqrt{\frac{t\tilde{\sigma}_t^2\eta^2+1}{t^2\eta^2}\log\left(\frac{t\tilde{\sigma}_t^2\eta^2+1}{\alpha^2}\right) + o(V_t\log V_t/t^2)} \\
&= \sqrt{\frac{t\tilde{\sigma}_t^2\eta^2+1}{t^2\eta^2}\log\left(\frac{t\tilde{\sigma}_t^2\eta^2+1}{\alpha^2}\right)} + o(\sqrt{V_t\log V_t}/t),
\end{aligned}$$

where the last line follows because $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ for $a, b \geq 0$. This formally shows how our confidence sequence in Theorem 4.1 is a valid $(1-\alpha)$−asymptotic confidence sequence for μ_t with approximation rate $o(\sqrt{V_t\log V_t}/t)$ given that our variance estimator is strongly consistent.

However, Theorem 2.1 only tells us that $\tilde{\sigma}_t^2$ is conditionally unbiased for $\bar{\sigma}_t^2$. To establish the consistency result, we again use a version of strong law of large numbers for martingale sequence difference. We denote $U_t := \tilde{\sigma}_t^2 - \bar{\sigma}_t^2$ and U_t is a martingale sequence difference with respect to the filtration $\mathcal{F}_{t-1}$. Our goal is to show that $U_t \to 0$ almost surely. First, we define $\tilde{U}_t := \sum_{i=1}^t U_i/i$. This quantity is bounded, a martingale-sequence difference, and also converges almost-surely because U_t satisfies the desired uniform integrability by Assumption 2.2. Finally applying Kronecker's Lemma on $\tilde{U}_t$ we get that $U_t \to 0$ almost surely.

F Optimizing and choosing η parameter

In this section, we show in detail how an analyst can choose η to optimize the confidence sequence width for a desired specific time t^*. The derivations are nearly identical to those presented in (Waudby-Smith et al., 2021), but we repeat them here for completeness.

Our confidence width presented in all the theorems have the following structure

$$B_t(\alpha) := \sqrt{\frac{2(t\eta^2+1)}{t^2\eta^2}\log\left(\frac{\sqrt{t\eta^2+1}}{\alpha}\right)},$$

where we have omitted the variance terms and instead substituted each $\hat{\sigma}_i = 1$ since we want η to be data-independent.[8] Then,

$$\operatorname*{argmin}_{\eta>0} B_t(\alpha) = \sqrt{\operatorname*{argmin}_{x>0} f(x)},$$

$$\text{where} \qquad f(x) := \frac{tx+1}{t^2 x} \log\left(\frac{tx+1}{\alpha^2}\right), \quad x := \eta^2.$$

Furthermore, $\lim_{x\to 0} f(x) = \lim_{x\to\infty} f(x)$ and thus if we can find the critical point by finding a solution for $\partial f / \partial x = 0$, then this must be the unique minimum.

Therefore, we have that

$$\frac{\partial f}{\partial x} = -\frac{1}{t^2 x^2} \log\left(\frac{tx+1}{\alpha^2}\right) + \frac{1}{tx}.$$

Setting the above to zero, we obtain

$$-\alpha^2 \exp(1) = -(tx+1)\exp(-(tx+1)).$$

Therefore, we have that the solution is $-(tx+1) = W_{-1}(-\alpha^2 \exp(1))$, where W_{-1} is the lower branch of the Lambert W function. The solution only exists if

$$-\alpha^2 \exp(1) \geq -\exp(1),$$

or equivalently if $\alpha^2 \leq 1$, which is always true for any $\alpha \in [0, 1]$. Therefore, we have that

$$\operatorname*{argmin}_{\eta>0} B_t(\alpha) = \sqrt{\frac{-W_{-1}(-\alpha^2 \exp(1)) - 1}{t^*}}.$$

[8] Consequently, we are not formally optimizing η for the actual confidence width, but η can still be conceptually interpreted as minimizing the confidence sequence width at a desired time t^* (See Appendix C.3 in (Waudby-Smith et al., 2021) for more details).

Appendix References

References

Fan, Xiequan, Ion Grama and Quansheng Liu. 2015a. "Exponential inequalities for martingales with applications." Electronic Journal of Probability 20:1 – 22.
URL: *https://doi.org/10.1214/EJP.v20-3496*

Fan, Xiequan, Ion Grama and Quansheng Liu. 2015b. "Exponential inequalities for martingales with applications." Electronic Journal of Probability 20:1 – 22.
URL: *https://doi.org/10.1214/EJP.v20-3496*

Howard, Steven R., Aaditya Ramdas, Jon D. McAuliffe and Jasjeet S. Sekhon. 2020. "Time-uniform, nonparametric, nonasymptotic confidence sequences." The Annals of Statistics.

Maurer, Andreas and Massimiliano Pontil. 2009. Empirical Bernstein Bounds and Sample-Variance Penalization. In Annual Conference Computational Learning Theory.

Ramdas, Aaditya, Johannes Ruf, Martin Larsson and Wouter M. Koolen. 2020. "Admissible anytime-valid sequential inference must rely on nonnegative martingales." arXiv: Statistics Theory.

Robbins, Herbert. 1970. "Statistical Methods Related to the Law of the Iterated Logarithm." The Annals of Mathematical Statistics 41:1397 – 1409.
URL: *https://doi.org/10.1214/aoms/1177696786*

Russo, Daniel. 2020. "Simple Bayesian Algorithms for Best-Arm Identification." Operations Research 68:1625–1647.
URL: *https://doi.org/10.1287/opre.2019.1911*

Slivkins, Aleksandrs. 2019. "Introduction to Multi-Armed Bandits." Foundations and Trends® in Machine Learning 12:1–286.
URL: *http://dx.doi.org/10.1561/2200000068*

Strassen, Volker. 1967. Almost sure behavior of sums of independent random variables and martingales.

Sutton, Richard and Andrew Barto. 2018. Reinforcement learning: an introduction. Adaptive Computation and Machine Learning. MIT Press.

Thompson, William R. 1933. "ON THE LIKELIHOOD THAT ONE UNKNOWN PROBABILITY EXCEEDS ANOTHER IN VIEW OF THE EVIDENCE OF TWO SAMPLES." Biometrika 25:285–294.
URL: *https://doi.org/10.1093/biomet/25.3-4.285*

Waudby-Smith, Ian, David Arbour, Ritwik Sinha, Edward H. Kennedy and Aaditya Ramdas. 2021. "Time-uniform central limit theory with applications to anytime-valid causal inference.".

Waudby-Smith, Ian, Lili Wu, Aaditya Ramdas, Nikos Karampatziakis and Paul Mineiro. 2022. "Anytime-valid off-policy inference for contextual bandits.".
URL: *https://arxiv.org/abs/2210.10768*

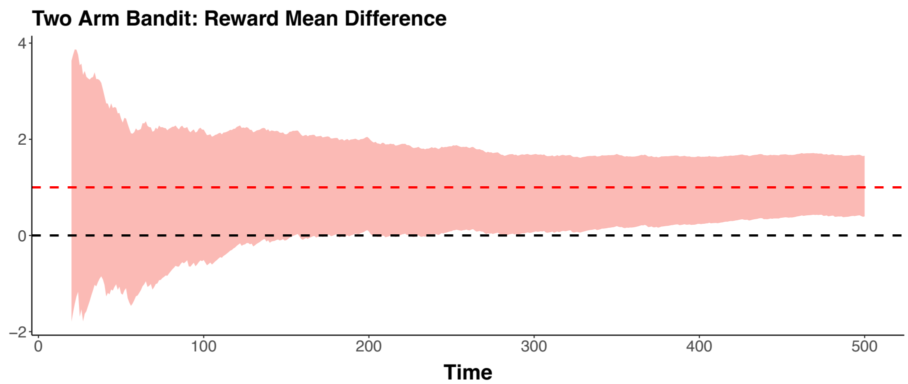

Figure 7: Two Arm Bandit (Example A.1). The red contours show the lower and upper confidence sequence as the agent pulls each arm adaptively at $\alpha = 0.05$ using Lemma 4.2 . The horizontal red dotted line represent the true mean difference of the rewards and the black horizontal dotted line represents the zero (null) line.

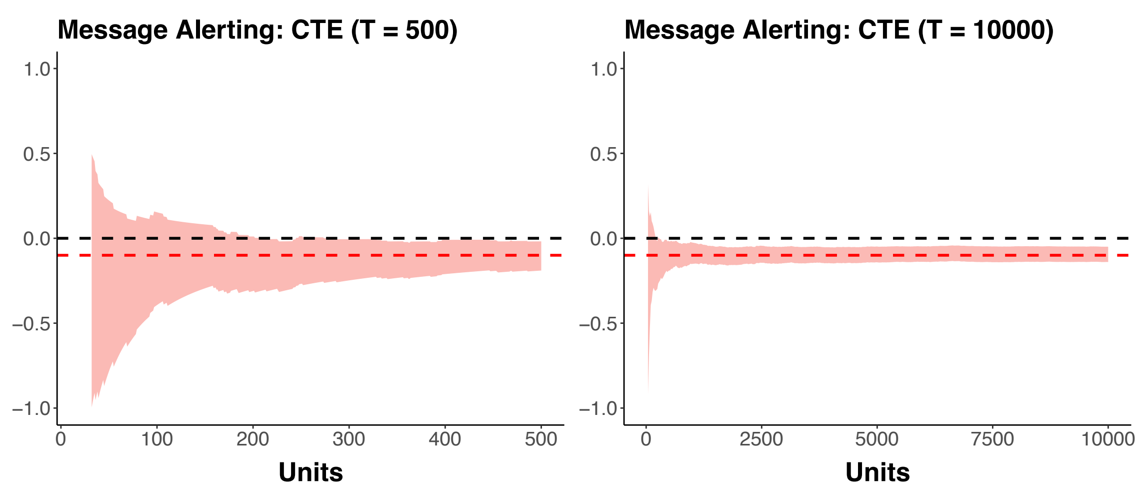

Figure 8: Online Experimentation (Example C.1). The red contours show the lower and upper confidence sequence as units enter in the experiment at $\alpha = 0.05$ using Theorem C.1. The left and right panel show the confidence sequences for $T = 500, 10000$ units, respectively. The horizontal red dotted line represent the true treatment effect of $\tau_n = 0.1$ and the black horizontal dotted line represents the zero (null) line.