Beyond Central Limit Theorem for Higher Order Inference in Batched Bandits

Anonymous Author(s)
Affiliation
Address
email

Abstract

Adaptive experiments have been gaining traction in a variety of domains, which stimulates a growing literature focusing on post-experimental statistical inference on data collected from such designs. Prior work constructs confidence intervals mainly based on two types of methods: (i) martingale concentration inequalities and (ii) asymptotic approximation to distribution of test statistics; this work contributes to the second kind. The current asymptotic approximation methods however mostly rely on first-order limit theorems, which can have a slow convergence rate in a data-poor regime. Besides, established results often rely on conditions that noises are well-behaved, which can be problematic when the real-world instances are heavy-tailed or asymmetric. In this paper, we propose the first higher-order asymptotic expansion formula for inference on adaptively collected data, which generalizes normal approximation to the distribution of standard test statistics. Our theorem relaxes assumptions on the noise distribution and benefits a fast convergence rate to accommodate small sample sizes. We complement our results by promising empirical performances in simulations.

1 Introduction

Adaptive experimental designs such as bandit algorithms have received increasing popularity in many applications[16, 2, 1]. Instead of fixing a randomization rule, the experimenter progressively updates the data collection mechanism in response to past observations, so as to reduce experimental costs or optimize sample efficiency to test hypothesis. With the increasing availability of data collected from such designs, we would like to understand the best practice to conduct post-experimental statistical inference. Specifically, we seek to evaluate alternative treatment assignment policies using data collected from batched multi-armed bandit algorithms and construct confidence intervals around the estimations. We would like to do so in a data-efficient and robust way and accommondate regimes where the sample size may be small, and the noise may be heavy-tailed or asymmetric.

There are two types of approaches of inference on adaptive data in the current literature. The first kind is based on martingale concentration inequalities that result in valid confidence intervals for arbitrary stopping times, but such methods are often overly conservative and lose statistical power especially when sample size is small. [12, 13, 9, 17]. Alternatively, asymptotic approximation-based methods construct estimators that have desirable limiting distributions, which are then translated to asymptotically valid confidence intervals [8, 20, 10, 21]. We view our work as complimentary to this line of research. Closest to our setting of batched multi-armed bandits, [20] first construct a studentized statistic per batch and then aggregate them to obtain an estimator with asymptotically normal guarantees, which retains even with nonstationarity.

Despite the growing research efforts, the current asymptotic approximation-based methods for post-experimental inference lack two things:

Submitted to the NeurIPS Causal ML for Real-World Impact Workshop (CML4Impact 2022). Do not distribute.

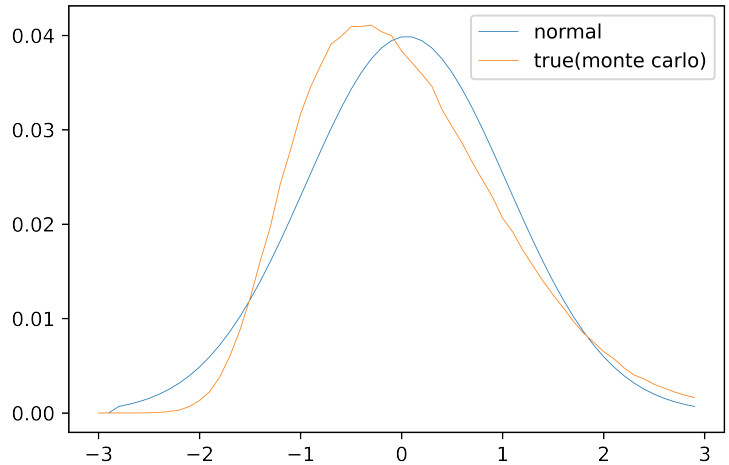

Figure 1: The monte carlo distribution of the test statistic based on OLS is far from being normal due to the asymmetric and heavy-tailed noise.

1. Robustness to distributional assumptions. Prior work has mainly relied on assumptions on sub-gaussian or symmetric noises, which is however problematic since such desirable properties may not be realized in many real-world applications. For example, treatment effect heterogeneity or truncation induces asymmetry in the noises. Thus methods predicated upon such assumptions might not yield correct statistical properties in cases of heavy-tailed or asymmetric noises.

2. Sample size considerations. The current asymptotic approximation methods for adaptive inference all rely on first-order limit theorems, a.k.a. central limit theorem. This convergence rate is slow in a data-poor regime, which is typical if we want to evaluate sub-optimal arms using bandit data. Bandit algorithms are often geared towards minimizing cumulative regret, which leads to meager data collection for arms that do not perform well during the first part of the experiment. This issue exacerbates as the number of arms increases.

To provide intuition, Example 1 shows that batched OLS [20] generates a test statistic that is far from being normal with limited sample size and asymmetric noise.

Example 1. *Consider a two-batch multi-armed bandit experiment run under a ϵ-greedy algorithm, where the noise is distributed as standardized $\Gamma(2, 1)$ with mean 0 and variance 1. Figure 1 presents the monte carlo distribution of the test statistic based on batched OLS [20], which is shown to be far from being normal due to the asymmetric and heavy-tailed noise.*

How should we tackle these problems? The confidence sequence approach has an appeal of time-uniform valid inference, but past simulation studies suggest that they may be conservative with small sample size[20, 8]. This paper provides an alternative method: if the first-order approximation is insufficient, why not resort to higher orders? With given sample size, our method aims to capture higher-order departure of the test statistic from being normally distributed and translate it into a form of an asymptotic expansion density. This asymptotic expansion formula formally guarantees rates of convergence of $o(n^{p-2})$, where p is a user-specified integer. This asymptotic expansion is used to provide higher-order adjusted z-scores for test statistics, which should be instance-dependent and with respect to parameters of the experiment such as sample size. This paper provides the first attempt on higher-order approximation in the context of adaptive experiments. Such attempt has been established in [18] when data is i.i.d. collected; our paper extends these techniques in [18] to the context of batch-wise adaptive data collection and applies it iteratively for each batch. Besides, we also allow for possibly discontinuous policy functions, which generalizes the established results for continuous policy class in [18].

The rest of the paper is organized as follows. Section 2 formulates the problem and provides preliminaries. Section 3 provides our main result, where we estimate an arm based on an asymptotic

expansion density that generalizes the commonly-made normal approximation. Section 4 offers preliminary numerical experiments, which are extended in the Appendix B. Finally, Section 5 concludes.

Two caveats are worth mentioning here: (i) We will try to be careful about measurability conditions since such conditions are crucial for delicate higher-order theory. Also, classical literature on adaptive experiments has documented how measurability conditions can have nontrivial practical considerations [5]. (ii) Due to the page limit, we refer interesting readers to Appendix A for more elaboration and formal proofs and to Appendix B for further numerical results.

2 Setup

2.1 Batched Bandit Framework

Suppose the agent interacts with n individuals at batch s, and the experiment has $\mathbb{S} = \{1, ..., S\}$ batches in total. At batch s, the i-th individual is identified with an element $j = (s, i_s)$. We write $\mathbb{J}^n = \{(s, i_s;\ i_s \in \mathbb{I}^n_s,\ s \in \mathbb{S}\}$, where $\mathbb{I}^n_s = \{1, ..., n_s\}$. Let $s(j) = s$ for each individual $j \in \mathbb{J}^n_s$. The batch size n is the parameter that drives the asymptotic theory we will develop.

Given a probability space $(\Omega, \mathcal{F}, P)$, for the individual $j \in \mathbb{J}^n_s$ the agent selects an action/treatment according to an arm assignment policy (to be shortly introduced), which is expressed by a $\overline{k}_s$-dimensional vector $A_j = (A_{j,k_s})_{k_s \in \mathcal{K}_s}$, where $\mathcal{K}_s$ is the action space of size k_s in batch s. Each entry A_{j,k_s} takes values in $\{1, 0\}$ and $\sum_{k_s \in \mathcal{K}_s} A_{j,k_s} = 1$.

The agent then observes a reward $R_j \in \mathbb{R}^{r_s}$ from the action A_j, which is written as follows,

$$R_j = A_j^\star \beta_{s(j)} + \epsilon_j \quad (j \in \mathbb{J}^n_s) \tag{1}$$

where $\beta_s \in \mathbb{R}^{\overline{k}_s}$ represents the underlying effect of the actions, $\dot{\epsilon}_j \in \mathbb{R}^{r_s}$ is the exogenous mean-zero noise that is sampled i.i.d. each time, and the star $\star$ denotes the matrix transpose[1].

For batch s, let $\mathbf{A}^n_s = (A_j;\ j \in \mathbb{J}^n_s)$ represent the action set and $\boldsymbol{\epsilon}^n_s = (\epsilon_j;\ j \in \mathbb{J}^n_s)$ represent the noise set. For $j \in \mathbb{J}^n_s$, let W^n_j be a $\mathsf{d}_s \times \mathsf{r}_s$ random matrix measurable with respect to $\sigma[\mathbf{A}^n_{s(j)}]$ Write $\mathbf{W}^n_s = (W^n_j)_{j \in \mathbb{J}^n_s}$. Consider a weighted sum

$$\mathbb{Z}^n_s = \sum_{j \in \mathbb{J}^n_s} W^n_j \epsilon_j. \tag{2}$$

This $\mathbb{Z}^n_s$ captures the most statistical quantities of interest in estimating the value of a single arm or arm difference for batch s. For example, using data collected only from batch s, the OLS estimator has the following form:

$$\mathbb{Z}^n_s = \sum_{j \in \mathbb{J}^n_s} W^n_j \epsilon_j = \mathrm{diag}((N^n_{s,1})^{-1/2} \sum_{j \in \mathbb{J}^n_s} A_{j,1}\epsilon_j, ..., (N^n_{s,\overline{k}_s})^{-1/2} \sum_{j \in \mathbb{J}^n_s} A_{j,\overline{k}_s}\epsilon_j),$$

where $N^n_{s,k}$ denotes the number of action k being selected in batch s.

We next model the bandit algorithm that determines the policy for each batch. Let $\mathcal{L}_s$ be a measurable space for $s \in \mathbb{S}$. We consider a $\sigma[\mathbf{A}^n_s]$-measurable random map $L^n_s : \Omega \to \mathcal{L}_s$ for every $(n, s) \in \mathbb{N} \times \mathbb{S}$. The variables L^n_{s-1} will be used for making a criterion for selection of an action in batch s. We write $\underline{L}^n_s = (L^n_1, ..., L^n_s)$ for $(L_s)_{s \in \mathbb{S}}$ [2]. In the batched bandits, how the agent updates the data collection policy c_s (the distribution of $\mathbf{A}^n_s$) from a measurable set $\mathfrak{C}_s$ is determined by the average effects of actions at batch $s - 1$. We represent the policy c_s as the the distribution $q\big((\underline{L}^n_{s-1}, \underline{\mathbb{Z}}^n_{s-1}), dc_s\big)$ over the action space.

Example 2 (ϵ-Greedy algorithm)**.** *The ϵ-Greedy algorithm having $\mathfrak{C}_2 = \{c_2^{(1)}, c_2^{(2)}\}$ with*

$$q\big((\underline{l}_1, \underline{z}_1), dc_2\big) = 1_{\{\mathtt{h}_1(\underline{l}_1, \underline{z}_1) \geq 0\}} \delta_{c_2^{(1)}}(dc_2) + 1_{\{\mathtt{h}_1(\underline{l}_1, \underline{z}_1) < 0\}} \delta_{c_2^{(2)}}(dc_2) \tag{3}$$

for some function $\mathtt{h}_1$.

[1] Extensions for considering different sets of actions and different distributions of ϵ_j for different batchs, or even incorporating contexts is straightforward though we adopted the present setting for notational simplicity.

[2] This operation by the underline $\underline{\cdot}_s$ will apply to other vectors

Example 3 (Thomspon sampling)**.** *A Thompson sampling algorithm is realized as*

$$q\big((\underline{l}_1, \underline{z}_1), dc_2\big) = 1_{\{\mathtt{h}_1(\underline{l}_1,\underline{z}_1)\le a_1\}}\delta_{c_2^{(1)}}(dc_2) + 1_{\{\mathtt{h}_1(\underline{l}_1,\underline{z}_1)> a_2\}}\delta_{c_2^{(2)}}(dc_2) + 1_{\{a_1<\mathtt{h}_1(\underline{l}_1,\underline{z}_1)\le a_2\}}\delta_{C_2(l_1,z_1)}(dc_2) \tag{4}$$

with some constants a_1, a_2 expressing the clipping constraint commonly assumed in batched bandits[20] and some $\mathfrak{C}_2$-valued function C_2 of (l_1, z_1).

Finally, we denote $\mathcal{G}_s^n = \sigma\big[L_{s'}^n, \mathbb{Z}_{s'}^n;\ s' \le s\big]$, and let $\mathcal{L} = \Pi_{s\in\mathbb{S}}\mathcal{L}_s$ and $\mathsf{d} = \sum_{s\in\mathbb{S}} \mathsf{d}_s$. We assume the below condition holds unless stated otherwise:

$$\boldsymbol{\epsilon}_s^n \amalg \left(\mathcal{G}_{s-1}^n \vee \sigma[\mathbf{A}_s^n]\right) \quad (s \in \mathbb{S}). \tag{5}$$

For notation convenience, we denote the conditional expectation $E[\,\cdot\,|\mathcal{G}_{s-1}^n]$ as $E_{s-1}[\,\cdot\,]$ and the conditional expectation $E\big[\ \cdot\ |\mathcal{G}_{s-1}^n \vee \sigma[\mathbf{A}_s^n]\big]$ as $E_{s-1,\mathbf{A}_s^n}[\,\cdot\,]$. For $\mathbf{w}_s^n = (w_j^n)_{j\in\mathbb{J}_s^n} \in \mathbb{R}^{\mathsf{d}_s \mathsf{r}_s n_s}$, let $P_s^n(\mathbf{w}_s^n, dz_s) = P^{\sum_{j\in\mathbb{J}_s^n} w_j^n \epsilon_j}(dz_s)$ represents the distribution of $\sum_{j\in\mathbb{J}_s^n} w_j^n \epsilon_j$ [3] Since L_s^n and $\mathbf{W}_s^n$ are $\sigma[\mathbf{A}_s^n]$-measurable, the assumption (18) yields a representation of a regular conditional distribution of $(L_s^n, \mathbf{W}_s^n)$ as

$$P^{(L_s^n,\mathbf{W}_s^n)}(dl_s, d\mathbf{w}_s^n|\mathcal{G}_{s-1}^n) = \int_{\mathfrak{C}_s} q\big((\underline{L}_{s-1}^n, \underline{\mathbb{Z}}_{s-1}^n), dc_s)\nu_{c_s}^n(dl_s, d\mathbf{w}_s^n) \tag{6}$$

for the probability distributions $\nu_{c_s}^n = \eta_{c_s}^{(L_s^n,\mathbf{W}_s^n)}$.

2.2 Off-policy Evaluation with Backwards Induction

When the experiment finishes, we want to conduct hypothesis testing like the value of the arm $\beta_{s(j)}$ for scientific discovery. We can generically define hypothesis testing problems mathematically as follows. Given a measurable function $f : \mathcal{L} \times \mathbb{R}^{\mathsf{m}} \to \mathbb{R}$, where $\mathsf{m} = \sum_{s\in\mathbb{S}} \mathsf{m}_s$, our estimand is the value of the expectation $\overline{\mathbf{E}}_n = E\big[f\big(\underline{L}_S^n, \underline{\mathbb{Y}}_S^n\big)\big]$. For standard hypotheses and confidence intervals, f is usually an indicator function for whether the test statistic is smaller than some particular threshold.

If we know the true underlying density $P_s^n(\mathbf{w}_s^n, dz_s)$, we can iteratively calculate the desired expectation from a simple exercise analogous to dynamic programming, with a slight modification with a truncation $1_{\{L_s^n\in\mathcal{A}_s^n\}}$, where $\mathcal{A}_s^n$ is a measurable set of $\mathcal{L}_s$. If the clipping constraint holds, $\mathbf{E}_n$ almost equals to $\bar{\mathbf{E}}_n := E\left[f\big(\underline{L}_S^n, \underline{\mathbb{Y}}_S^n\big)\prod_{s=1}^S 1_{\{L_s^n\in\mathcal{A}_s^n\}}\right]$ [4]. Theorem3 formally proves this with the aid of the following Lemma.

Lemma 1. *For $s \in \mathbb{S}$,*

$$E_{s-1}\left[f_s^n\big(\underline{L}_s^n,\ \underline{\mathbb{Z}}_s^n\big)\prod_{s'=1}^{s} 1_{\{L_{s'}^n\in\mathcal{A}_{s',c}^n\}}\right] = f_{s-1}^n(\underline{L}_{s-1}^n,\ \underline{\mathbb{Z}}_{s-1}^n)\prod_{s'=1}^{s-1} 1_{\{L_{s'}^n\in\mathcal{A}_{s',c}^n\}} \quad a.s. \tag{7}$$

Here the product $\prod_{s'=1}^0$ reads 1.

Theorem 1. *Suppose that $f\big(\underline{L}_S^n, \underline{\mathbb{Z}}_S^n\big)$ is integrable. Then*

(a) $\mathbf{E}_n = f_0^n$, *where* $f_0^n := \int f_{1,c}^n\big(l_1,\ z_1\big)P_1^n(\mathbf{w}_1^n, dz_1)1_{\{l_1\in\mathcal{A}_1^n\}}\nu_{c_1}^n(dl_1, d\mathbf{w}_1^n)$,

$$f_S^n(\underline{l}_S, \underline{z}_S) = f\big(\underline{l}_S, (Y_s^n(l_s, z_s))_{s\le S}\big)$$

$$f_{s-1}^n = f_{s-1}^n(\underline{l}_{s-1}, \underline{z}_{s-1})$$

$$= E\left[\int_{\mathfrak{C}_s}\int_{\mathcal{L}_s\times(\mathbb{R}^{\mathsf{d}_s}\otimes\mathbb{R}^{\mathsf{r}_s})} f_s^n\big(\underline{l}_{s-1}, l_t, \underline{z}_{s-1}, \textstyle\sum_{j\in\mathbb{J}_s^n} w_j^n\epsilon_j\big)1_{\{l_s\in\mathcal{A}_s^n\}}\nu_{c_s}^n(dl_s, d\mathbf{w}_s^n)q\big((\underline{l}_{s-1}, \underline{z}_{s-1}), dc_s\big)\right]$$

for $s \in \mathbb{S}$.

[3] In other words, $P_s^n(\mathbf{W}_s^n, dz_s)$ is a regular conditional probability of $\mathbb{Z}_s^n$

[4] The choice of $\mathcal{A}_{s,c}^n$ is important so that it will determine the accuracy of approximation $\mathbf{E}_n$ to $\overline{\mathbf{E}}_n$, but the clipping condition that is placed suffices for this condition. Intuitively, when there is too little sample size, there is no possibility of asymptotics.

(b) $\overline{\mathbf{E}}_n \;=\; E\big[f\big(\underline{L}_S^n, \underline{\mathbb{Y}}_S^n\big)\big] \;=\; \mathbf{E}_n + \rho_n$ *with* $|\rho_n| \le E\Big[\big|f\big(\underline{L}_S^n, \underline{\mathbb{Z}}_S^n\big)\big| \sum_{s\in\mathbb{S}} 1_{\{L_s^n \notin \mathcal{A}_s^n\}}\Big]$.

Theorem 3 on its own does not have much practical relevance, since the true distribution $P_s^n(\mathbf{w}_s^n, dz_s)$ is unknown. Therefore past research on off policy evaluation have aimed to approximate it by some random signed measure $\Psi^n_{s,p,\mathbf{w}_s^n}$on $\mathbb{R}^d$depending on $\mathbf{w}_s^n$. All past attempts on asymptotic approximation claimed that the gaussian density was sufficient. We claimed that can be problematic especially when the sample size is small and the noise is not well behaved.

3 Main Result: Asymptotic Expansion for Batched Bandits

To assess the accuracy of the approximation, we need to estimate the difference between f_0 and $\widehat{f}_0$, where $\widehat{f}_0$ is the approximation to the f_0 by replacing $P_s^n(\mathbf{w}_s^n, dz_s)$ with $\Psi^n_{s,p,\mathbf{w}_s^n}$ for each batch; here f_0^n was defined as $f_0^n = \int f^n_{1,c}\big(l_1,\ z_1\big) P_1^n(\mathbf{w}_1^n, dz_1) 1_{\{l_1\in\mathcal{A}_1^n\}} \nu^n_{c_1}(dl_1, d\mathbf{w}_1^n)$.

We shall only present the results with assumptions stated verbally due to the page limit and defer the mathematical formulation of the assumptions to Appendix A. The conditions we assume include (i) conditional i.i.d. noise ϵ that hold in most sequential experiments, (ii) regularity conditions such as measurability or existence of moments, and (iii) clipping conditions that are usually implicitly or explicitly assumed in the past literature [18]. The only condition that uniquely arises in the higher-order case is the well-known Cramér condition as follows, for any batch $s \in \mathbb{S}$, there exists a B_0 such that

$$\inf_{u\in\mathbb{R}^{d_s}:|u|\ge B_0} |E_{\mathcal{C}^s}[\exp(iu\cdot\epsilon_{(s,1)})]| < 1,$$

which places regularity on the characteristic function of the noise distribution and validates the Fourier inversion to get the asymptotic expansion density. This condition generally holds except for lattice distributions that we leave for future work.

Theorem 2 (Informal)**.** *Suppose the conditions stated above hold. There exists a constant M^* such that $\mathbf{E}_n = f_0^n$ admits the following estimates:*

$$|f_0^n - \widehat{f}_0^n| \quad\le\quad M^*\Big(n^{-(\mathsf{p}-2+\delta_1)/2}\sum_{s\in\mathbb{S}}\prod_{i=2}^{s} U_i^n\Big), \tag{8}$$

for all $f \in \widehat{\mathcal{D}}(\mathbf{M},\gamma)$ and a class of measurable functions with at most polynomial growth $\mathbf{M} = (M_s)_{s\in\mathbb{S}}$, where

$$U_{s+1}^n \quad=\quad \int (1+|z_s|^\gamma)|\Psi^n_{s,\mathsf{p},\mathbf{w}_s^n}|(dz_s) 1_{\{l_s\in\mathcal{A}_s^n\}} \nu^n_{c_s}(dl_s, d\mathbf{w}_s^n) q\big((\underline{l}_{s-1}, \underline{z}_{s-1}), dc_s\big). \tag{9}$$

for $s = 1, ..., S-1$.

Remark 1. *Theorem 34 suggests using $\widehat{f}^n_{0,c}$ to approximate $\mathbf{E}_{n,c}$. Usually U_{s+1}^n is uniformly bounded in n or grows as slowly as to be controlled by the factor $n^{-\delta_1/2}$ in the error bound of (74).*

We also present an explicit form of the asymptotic expansion density that can be used to calculate the higher order corrected Z-score for BOLS estimator[20] up to order p (which is a user-specified integer) when the variance is known. Appendix A shows the expression when the variance is undisclosed.

Proposition 1. *The function $d\Psi_{n_s,p,\mathcal{C}^s}/dz(z)$ ($z = (z_1, ..., z_{\overline{k}_s})$) is*

$$\prod_{k\in\mathcal{K}_s}\phi(z_k;\lambda_{n_s,2,\mathcal{C}^s,k}) \times \prod_{k\in\mathcal{K}_s}\Big\{1 + \frac{1}{6}n_s^{-1/2}\lambda_{n_s,3,\mathcal{C}^s,k}h_3(z_k;\lambda_{n_s,2,\mathcal{C}^s,k})$$
$$+ n_s^{-1}\Big(\frac{1}{24}\lambda_{n,k,4}h_4(z_k;\lambda_{n,k,2}) + \frac{1}{72}\lambda^2_{n_s,k,3}h_6(z_k;\lambda_{n_s,k,2})\Big) + \cdots\Big\}, \tag{10}$$

with the summation taken up to order $n_s^{-(p-2)/2}$, where $\lambda_{n_s,r,\mathcal{C}^s,k} = n_s^{(r-2)/2}\cdot (N^n_{s,k}/n_s)^{-(r-2)/2}\,\kappa_{r,\mathcal{C}^s}(\epsilon_{(s,1)})(\kappa_{r,\mathcal{C}^s}(\epsilon_{(s,1)})$ is the rth order $\mathcal{C}^s$conditional cumulant of $\epsilon_{(s,1)}$, and $h_r(z;\Sigma) = (-1)^r\phi(z;0,\Sigma)^{-1}\partial_z^r\phi(z;0,\Sigma)$ is the r-th order Hermite polynomial.

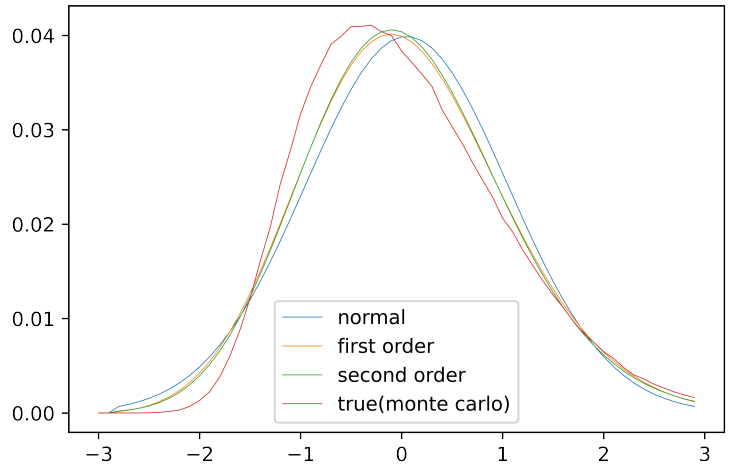

Figure 2: Higher-order asymptotic expansion provides better finite-sample approximation of the test statistics as compared to normal distribution.

Remark 2. *Note that all elements of this density are known and calculable for batched bandits. Even if we don't know the distribution and hence the moments of $\epsilon_{(s,1)}$, we derived an asymptotic expansion that is valid for first order asymptotic expansion formula that is valid under plugged in moments. See appendix B for the derivation and the general formula.*

4 Simulations

In the main paper, we only provide the numerical results for the setting described in Example 1. As figure 2 shows, asymptotic expansion has clear gains over the normal distribution in approximating the true underlying density. While the second order asymptotic expansion seems to have better approximation in most regions of the support, the difference may not be that large. Expanding to higher order may further capture the deviations from normality. We defer extensive experiments that vary sample size and the regularity of the noise distribution to Appendix B.

5 Discussions

In this paper, we have introduced, developed, and justified the first theory for higher-order statistical inference specialized for data collected by batched bandit algorithms. Interesting future work includes extending our higher order theory to non-batched case and doing power analysis based on higher order theory. Having better type-1 error control and powered experiments thanks to higher order inference has a potential to extend the limits on the experimental design, enabling more aggressive algorithms or reducing the sample sizes necessary.

References

[1] Supplement for: Adaptive treatment assignment in experiments for policy choice.

[2] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire. The nonstochastic multiarmed bandit problem. *SIAM journal on computing*, 32(1):48–77, 2002.

[3] R. N. Bhattacharya and J. K. Ghosh. On the validity of the formal Edgeworth expansion. *Ann. Statist.*, 6(2):434–451, 1978.

[4] R. N. Bhattacharya and R. R. Rao. *Normal approximation and asymptotic expansions*, volume 64. SIAM, 2010.

[5] W. Brannath, G. Gutjahr, and P. Bauer. Probabilistic foundation of confirmatory adaptive designs. *Journal of the American Statistical Association*, 107(498):824–832, 2012.

[6] Y. Deshpande, L. Mackey, V. Syrgkanis, and M. Taddy. Accurate inference for adaptive linear models. *PMLR*, 2018.

[7] M. Dudík, J. Langford, and L. Li. Doubly robust policy evaluation and learning. *arXiv preprint arXiv:1103.4601*, 2011.

[8] V. Hadad, D. A. Hirshberg, R. Zhan, S. Wager, and S. Athey. Confidence intervals for policy evaluation in adaptive experiments. *PNAS*, 2021.

[9] S. R. Howard, A. Ramdas, J. Mcauliffe, and J. Sekhon. Time-uniform, nonparametric, nonasymptotic, confidence sequences. *Annals of Statistics*, 2021.

[10] G. Lewis and V. Syrgkanis. Double/debiased machine learning for dynamic treatment effects via g-estimation. *arXiv preprint arXiv:2002.07285*, 2020.

[11] A. R. Luedtke and M. J. v. d. Laan. Parametric-rate inference for one-sided differentiable parameters. *Journal of the American Statistical Association*, 113(522):780–788, 2018.

[12] A. Ramdas, J. Ruf, M. Larsson, and W. Koolen. Admissible anytime-valid sequential inference must rely on nonnegative martingales. *arXiv preprint arXiv:2009.03167*, 2020.

[13] H. Robbins. Statistical methods related to the law of the iterated logarithm. *The Annals of Mathematical Statistics*, 41(5):1397–1409, 1970.

[14] Y. Sakamoto and N. Yoshida. Asymptotic expansion formulas for functionals of ϵ-Markov processes with a mixing property. *Ann. Inst. Statist. Math.*, 56(3):545–597, 2004.

[15] Y. Sakamoto and N. Yoshida. Third-order asymptotic expansion of M-estimators for diffusion processes. *Ann. Inst. Statist. Math.*, 61(3):629–661, 2009.

[16] M. J. Van Der Laan and S. D. Lendle. Online targeted learning. 2014.

[17] I. Waudby-Smith and A. Ramdas. Estimating means of bounded random variables by betting. *arXiv preprint arXiv:2010.09686*, 2020.

[18] N. Yoshida. Partial mixing and Edgeworth expansion. *Probability Theory and Related Fields*, 129(4):559–624, 2004.

[19] N. Yoshida. *Suritoukeigaku (in Japanese)*, volume 21 of *Sugaku no Kangaekata*. Asakura Shoten, 2006.

[20] K. W. Zhang, L. Janson, and S. A. Murphy. Inference for batched bandits. *arXiv preprint arXiv:2002.03217*, 2020.

[21] K. W. Zhang, L. Janson, and S. A. Murphy. Statistical inference after adaptive sampling in non-markovian environments. *arXiv preprint arXiv:2202.07098*, 2022.

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes]
   (c) Did you discuss any potential negative societal impacts of your work? [Yes]
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes]
   (b) Did you include complete proofs of all theoretical results? [Yes]
3. If you ran experiments...

(a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes]

(b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes]

(c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [No] In the main section, we assume variance known for the asymptotic expansion densities, so our method should calculate the formula exactly. We present the cases with estimated variance in the appendix, where we show standard errors and the number of monte carlo replicates.

(d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes]

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...

(a) If your work uses existing assets, did you cite the creators? [Yes]

(b) Did you mention the license of the assets? [Yes]

(c) Did you include any new assets either in the supplemental material or as a URL? [Yes]

(d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [Yes]

(e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [Yes]

5. If you used crowdsourcing or conducted research with human subjects...

(a) Did you include the full text of instructions given to participants and screenshots, if applicable? [Yes]

(b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [Yes]

(c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [Yes]

A Formal framework and proofs

As mentioned in the main body, formal assumptions and proofs, had to be omitted. We present the complete theory below. For readers only interested in the direct proofs of Lemma1, Theorem1, Theorem2, Proposition1, please refer to Lemma2, Theorem3, Theorem7, the explanation in C.0.1, respectively.

A.1 A model of bandit algorithms and a backward recursion formula

A.1.1 A model of bandit algorithms

Let $\mathbf{A}_s^n = (A_j;\ j \in \mathbb{J}_s^n)$ and $\boldsymbol{\epsilon}_s^n = (\epsilon_j;\ j \in \mathbb{J}_s^n)$ for $s \in \mathbb{S}$, where ϵ_j is an r_s-dimensional random vector. [E.g., $\epsilon_j = (\dot{\epsilon}_j, \dot{\epsilon}_j^2 - \sigma^2)^\star$.] For $j \in \mathbb{J}_s^n$, let W_j^n be a $\mathsf{d}_s \times \mathsf{r}_s$ random matrix measurable with respect to $\sigma[\mathbf{A}_{s(j)}^n]$ Write $\mathbf{W}_s^n = (W_j^n)_{j \in \mathbb{J}_s^n}$. Consider a weighted sum

$$\mathbb{Z}_s^n \quad = \quad \sum_{j \in \mathbb{J}_s^n} W_j^n \epsilon_j. \tag{11}$$

$\mathbb{Z}_n = \sum_{t \in \mathbb{T}} \mathbb{Z}_s^n$, i.e.,

$$\mathbb{Z}_n \quad = \quad \sum_{j \in \mathbb{J}^n} W_j^n \epsilon_j \tag{12}$$

Let $\mathcal{L}_s$ be a measurable space for $s \in \mathbb{S}$. We consider a $\sigma[\mathbf{A}_s^n]$-measurable random map $L_s^n : \Omega \to \mathcal{L}_s$ for every $(n, s) \in \mathbb{N} \times \mathbb{S}$. The variables L_{s-1}^n will be used for making a criterion for selection of a treatment at batch s. We write $\underline{L}_s^n = (L_1^n, ..., L_s^n)$ for $(L_s)_{s \in \mathbb{S}}$. This operation by the underline $\underline{\cdot}_s$ will apply to other vectors. In the batched bandits, the strategy (the distribute of $\mathbf{A}_s^n$) is determined by information up to batch $s-1$, e.g., the average effects of actions at batch $s-1$. Suppose that an distribution labeled by c_s is chosen from a measurable set $\mathfrak{C}_s$ for each $s \in \mathbb{T}$ following the

probability distribution $q\big((\underline{L}^n_{s-1}, \underline{\mathbb{Z}}^n_{s-1}), dc_s\big)$ on $\mathfrak{C}_s$. We assume that $\#\mathfrak{C}_1 = 1$ and $q(\emptyset, dc_1)$ is the Dirac measure on the single point $c_1 \in \mathfrak{C}_1$. , and that the choice c_s is equivalent to

$$(\underline{L}^n_{s-1}, \underline{\mathbb{Z}}^n_{s-1}) \in B(c_s)$$

for some measurable set $B(c_s)$ in $\big(\Pi^s_{s'=1}\mathcal{L}_{s'}\big) \times \big(\Pi^s_{s'=1}\mathbb{R}^{\mathsf{d}_{s'}}\big)$. Since some c_s is chosen at batch s,

$$\sum_{c_s \in \mathfrak{C}_s} 1_{\big\{(\underline{L}^n_{s-1}, \underline{\mathbb{Z}}^n_{s-1}) \in B(c_s)\big\}} \quad = \quad 1 \quad a.s. \tag{13}$$

From (13),

$$1_{\big\{(\underline{L}^n_{s-1}, \underline{\mathbb{Z}}^n_{s-1}) \in B(k)\big\}} 1_{\big\{(\underline{L}^n_{s-1}, \underline{\mathbb{Z}}^n_{s-1}) \in B(k')\big\}} \quad = \quad 1_{\{k=k'\}} 1_{\big\{(\underline{L}^n_{s-1}, \underline{\mathbb{Z}}^n_{s-1}) \in B(k)\big\}} \quad a.s. \tag{14}$$

for $k, k' \in \mathfrak{C}_s$. We assume $\#\mathfrak{C}_1 = 1$ and set $\big\{\omega \in \Omega;\ (\underline{L}^n_0, \underline{\mathbb{Z}}^n_0) \in B(c_1)\big\} = \Omega$ for convenience. In particular, (13) is satisfied for all $s \in \mathbb{S}$. Furthermore, we set $B(c_{S+1})$ as the whole space and $\#\mathfrak{C}_{S+1} = 1$, therefore $\big\{\omega \in \Omega;\ (L^n_S, \mathbb{Z}^n_S) \in B(c_{S+1})\big\} = \Omega$. In Section C.0.2, we will treat a linear combination of some basic statistics associated to $\mathbb{Z}^n_s$, with $\sigma[\mathbf{A}^n_s]$-measurable random weights incorporated into the variable L^n_s; see (91). So, maybe having two components, L^n_s plays two roles: defining assignment mechanism and integrating basic statistics. For each $(n, t) \in \mathbb{N} \times \mathbb{T}$, we consider a $\sigma[\mathbf{A}^n_t]$-measurable random variable $V^n_s : \Omega \to \mathcal{V}_t$ for a measurable space $\mathcal{V}_t$.

The history is recorded by the σ-fields $\mathcal{H}^n_s = \sigma[\mathbf{A}^n_{s'}, \boldsymbol{\epsilon}_{s'}\ s' \leq s]$ for $s \in \{0\} \cup \mathbb{S}$ with the trivial σ-field $\mathcal{H}^n_0$. An often used assumption is the independency between $\boldsymbol{\epsilon}^n_s$ and $\mathcal{H}^n_{s-1} \vee \sigma[\mathbf{A}^n_s]$, i.e.,

$$\boldsymbol{\epsilon}^n_s\big(\mathcal{H}^n_{s-1} \vee \sigma[\mathbf{A}^n_s]\big) \quad (s \in \mathbb{S}) \tag{15}$$

Condition (15) is standard one but we can relax this condition for our use. Let

$$\mathcal{G}^n_s = \sigma\big[L^n_{s'}, \mathbb{Z}^n_{s'};\ s' \leq s\big].$$

Instead of (15), we assume the condition

$$\boldsymbol{\epsilon}^n_s\big(\mathcal{G}^n_{s-1} \vee \sigma[\mathbf{A}^n_s]\big) \quad (s \in \mathbb{S}) \tag{16}$$

unless otherwise stated. The conditional expectation $E[\,\cdot\,|\mathcal{G}^n_{s-1}]$ is denoted by $E_{s-1}[\,\cdot\,]$, and the conditional expectation $E\big[\ \cdot\ |\mathcal{G}^n_{s-1} \vee \sigma[\mathbf{A}^n_s]\big]$ by $E_{s-1,\mathbf{A}^n_s}[\,\cdot\,]$.

For $\mathbf{w}^n_s = (w^n_j)_{j \in \mathbb{J}^n_s} \in \mathbb{R}^{\mathsf{d}_s \mathsf{r}_s n_s}$, let

$$P^n_s(\mathbf{w}^n_s, dz_s) \quad = \quad P^{\sum_{j \in \mathbb{J}^n_s} w^n_j \epsilon_j}(dz_s), \tag{17}$$

the distribution of $\sum_{j \in \mathbb{J}^n_s} w^n_j \epsilon_j$. In other words, $P^n_s(\mathbf{W}^n_s, dz_s)$ is a regular conditional probability of $\mathbb{Z}^n_s$ given $\mathbf{W}^n_s$ becuase of (16).

We assume

$$P^{\mathbf{A}^n_s}\big(da_s|\mathcal{G}^n_{s-1}\big) \quad = \quad \int_{\mathfrak{C}_s} q\big((\underline{L}^n_{s-1}, \underline{\mathbb{Z}}^n_{s-1}), dc_s\big)\eta_{c_s}(da_s) \tag{18}$$

for some probability distributions η_{c_s} $(c_s \in \mathfrak{C}_s)$ for a regular conditional distribution $P^{\mathbf{A}^n_s}\big(da_s|\mathcal{G}^n_{s-1}\big)$ of $\mathbf{A}^n_s$ given $\mathcal{G}^n_{s-1}$. Since L^n_s and $\mathbf{W}^n_s$ are $\sigma[\mathbf{A}^n_s]$-measurable, the assumption (18) yields a representation of a regular conditional distribution of $(L^n_s, \mathbf{W}^n_s)$ as

$$P^{(L^n_s, \mathbf{W}^n_s)}(dl_s, d\mathbf{w}^n_s|\mathcal{G}^n_{s-1}) \quad = \quad \int_{\mathfrak{C}_s} q\big((\underline{L}^n_{s-1}, \underline{\mathbb{Z}}^n_{s-1}), dc_s\big)\nu^n_{c_s}(dl_s, d\mathbf{w}^n_s) \tag{19}$$

for the probability distributions $\nu^n_{c_s} = \eta^{(L^n_s, \mathbf{W}^n_s)}_{c_s}$.

Let $\mathcal{L} = \Pi_{s \in \mathbb{S}}\mathcal{L}_s$. Following the already mentioned rule, we write $\underline{\mathbb{Z}}^n_s = \big(\mathbb{Z}^n_1, ..., \mathbb{Z}^n_s\big)$ and $\underline{z}_s = (z_1, ..., z_s)$ for $(z_1, ..., z_s) \in \mathbb{R}^{\mathsf{d}}$, $\mathsf{d} = \sum_{s \in \mathbb{S}} \mathsf{d}_s$. Consider an m_s-dimensional random variable

$$\mathbb{Y}^n_s \quad = \quad Y^n_s(L^n_s, \mathbb{Z}^n_s) \tag{20}$$

for a measurable map $Y^n_s : \mathcal{L}_s \times \mathbb{R}^{\mathsf{d}_s} \to \mathbb{R}^{\mathsf{m}_s}$. Equation (89) of Section C.0.2 gives an example of the function Y^n_s as the so-called Bhattacharya-Ghosh map, that was used in Bhattacharya and Ghosh [3] to validate the formal Edgeworth expansion for various statsitics. Given a measurable function,

$$f : \mathcal{L} \times \mathbb{R}^{\mathsf{m}} \to \mathbb{R}, \tag{21}$$

$\mathsf{m} = \sum_{s \in \mathbb{S}} \mathsf{m}_s$, the aim of this article is to compute the value of the expectation

$$\overline{\mathbf{E}}_n \quad = \quad E\big[f\big(\underline{L}^n_S, \underline{\mathbb{Y}}^n_S\big)\big], \tag{22}$$

suppose that $f\big(\underline{L}^n_S, \underline{\mathbb{Y}}^n_S\big)$ is integrable.

A.1.2 A backward recursion formula

Due to (13),

$$\overline{\mathbf{E}}_n = \sum_{c\in\mathfrak{C}} \overline{\mathbf{E}}_{n,c}, \tag{23}$$

where

$$\overline{\mathbf{E}}_{n,c} = E\Big[f\big(\underline{L}^n_S, \underline{\mathbb{Y}}^n_S\big) \prod_{s=1}^{S} 1_{\{(\underline{L}^n_s, \underline{\mathbb{Z}}^n_s)\in B(c_{s+1})\}} \Big] \tag{24}$$

and $c = (c_s)_{s\in\mathbb{S}} \in \mathfrak{C} = \prod_{s\in\mathbb{S}} \mathfrak{C}_s$. For $\overline{\mathbf{E}}_n$, it is sufficient to compute the value of $\overline{\mathbf{E}}_{n,c}$ for each $c \in \mathfrak{C}$. We will consider a modification of $\overline{\mathbf{E}}_n$ with a truncation $1_{\{L^n_s\in\mathcal{A}^n_s\}}$ with a measurable set $\mathcal{A}^n_s$ of $\mathcal{L}_s$. Let

$$\mathbf{E}_n = E\Big[f\big(\underline{L}^n_S, \underline{\mathbb{Y}}^n_S\big) \prod_{s=1}^{S} 1_{\{L^n_s\in\mathcal{A}^n_s\}} \Big]. \tag{25}$$

The choice of $\mathcal{A}^n_{s,c}$ is important so that it will determine the accuracy of approximation $\mathbf{E}_n$ to $\overline{\mathbf{E}}_n$. Let

$$f^n_S(\underline{l}_S, \underline{z}_S) = f\big(\underline{l}_S, (Y^n_s(l_s, z_s))_{s\le S}\big) \tag{26}$$

and define a function f^n_{s-1} by

$$\begin{aligned}
& f^n_{s-1}(\underline{l}_{s-1}, \underline{z}_{s-1}) \\
= \; & E\Big[\int_{\mathfrak{C}_s} \int_{\mathcal{L}_s\times(\mathbb{R}^{\mathsf{d}_s}\otimes\mathbb{R}^{\mathsf{r}_s})} f^n_s\big(\underline{l}_{s-1}, l_t, \underline{z}_{s-1}, \sum_{j\in\mathbb{J}^n_s} w^n_j \epsilon_j\big) 1_{\{l_s\in\mathcal{A}^n_s\}} \nu^n_{c_s}(dl_s, d\mathbf{w}^n_s) \\
& \qquad\qquad \times q\big((\underline{l}_{s-1}, \underline{z}_{s-1}), dc_s\big) \Big] \\
= \; & \int_{\mathfrak{C}_s} \int_{\mathcal{L}_s\times\mathbb{R}^{\mathsf{d}_s}} f^n_s\big(\underline{l}_{s-1}, l_t, \underline{z}_{s-1}, z_s\big) P^n_s(\mathbf{w}^n_s, dz_s) 1_{\{l_s\in\mathcal{A}^n_s\}} \nu^n_{c_s}(dl_s, d\mathbf{w}^n_s) \\
& \qquad\qquad \times q\big((\underline{l}_{s-1}, \underline{z}_{s-1}), dc_s\big)
\end{aligned} \tag{27}$$

for $s \in \mathbb{S}$. [Remark. In the previous formula, here we take sums w.r.t. c_{s+1} and c_s. The sum in c_{s+1} is included in f^n_s in the new formulation. So, the sum in c_s remains in the formula.] The function f^n_0 is a constant given by

$$f^n_0 = \int f^n_1\big(l_1, z_1\big) P^n_1(\mathbf{w}^n_1, dz_1) 1_{\{l_1\in\mathcal{A}^n_1\}} \nu^n_{c_1}(dl_1, d\mathbf{w}^n_1). \tag{28}$$

Now we have the following recursive formula.

Lemma 2. *For $s \in \mathbb{S}$,*

$$E_{s-1}\Big[f^n_s\big(\underline{L}^n_s, \underline{\mathbb{Z}}^n_s\big) \prod_{s'=1}^{s} 1_{\{L^n_{s'}\in\mathcal{A}^n_{s',c}\}} \Big] = f^n_{s-1}(\underline{L}^n_{s-1}, \underline{\mathbb{Z}}^n_{s-1}) \prod_{s'=1}^{s-1} 1_{\{L^n_{s'}\in\mathcal{A}^n_{s',c}\}} \quad a.s. \tag{29}$$

Here the product $\prod^0_{s'=1}$ reads 1.

Proof. We have

$$
\begin{aligned}
& E_{s-1}\Big[f_s^n\big(\underline{L}_s^n, \underline{\mathbb{Z}}_s^n\big)1_{\{L_s^n\in\mathcal{A}_s^n\}}\Big] \\
= \; & E_{s-1}\Big[E_{s-1,\mathbf{A}_s^n}\Big[f_s^n\big(\underline{L}_s^n, \underline{\mathbb{Z}}_{s-1}^n, \sum_{j\in\mathbb{J}_s^n} W_j^n\epsilon_j\big)\Big]1_{\{L_s^n\in\mathcal{A}_s^n\}}\Big] \\
= \; & E_{s-1}\Big[E\Big[f_s^n\big(\underline{l}_s, \underline{z}_{s-1}, \sum_{j\in\mathbb{J}_s^n} w_j^n\epsilon_j\big)\Big]\Big|_{\substack{\underline{z}_{s-1}=\underline{\mathbb{Z}}_{s-1}^n \\ \mathbf{w}_s^n=\mathbf{W}_s^n \\ \underline{l}_s=\underline{L}_s^n}} 1_{\{L_s^n\in\mathcal{A}_s^n\}}\Big] \qquad ((16)) \\
= \; & \int E\Big[f_s^n\big(\underline{l}_{s-1}, l_s, \underline{z}_{s-1}, \sum_{j\in\mathbb{J}_s^n} w_j^n\epsilon_j\big)\Big]1_{\{l_s\in\mathcal{A}_s^n\}} \\
& \qquad\qquad \times\nu_{c_s}^n(dl_s, d\mathbf{w}_s^n)q\big((\underline{l}_{s-1}, \underline{z}_{s-1}), dc_s\big)\Big|_{\substack{\underline{z}_{s-1}^n=\underline{\mathbb{Z}}_{s-1}^n \\ \underline{l}_{s-1}=\underline{L}_{s-1}^n}} \\
= \; & f_{s-1}^n(\underline{l}_{s-1}, \underline{z}_{s-1})\Big|_{\substack{\underline{z}_{s-1}^n=\underline{\mathbb{Z}}_{s-1}^n \\ \underline{l}_{s-1}=\underline{L}_{s-1}^n}} \qquad (30)
\end{aligned}
$$

by (19). Therefore we obtain (2).

$$
\begin{aligned}
= \; & E_{s-1}\Big[E\Big[f_{s,c}^n\big(\underline{l}_s, \underline{z}_{s-1}^n, \sum_{j\in\mathbb{J}_s^n} w_j^n\epsilon_j\big)1_{\left\{(l_s, \sum_{j\in\mathbb{J}_s^n} w_j^n\epsilon_j)\in B(c_{s+1})\right\}}\Big]\Big|_{\underline{z}_{s-1}^n=\underline{\mathbb{Z}}_{s-1}^n} 1_{\{l_s\in\mathcal{A}_{s,c}^n\}} \\
& \qquad\qquad \times\nu_{c_s}^n(dl_s, d\mathbf{w}_s^n)1_{\left\{(L_{s-1}^n, \mathbb{Z}_{s-1}^n)\in B(c_s)\right\}}\Big] \\
= \; & koko E_{s-1}\Big[\int f_{s,c}^n\big(\underline{l}_s, \underline{z}_{s-1}^n, z_s\big)P_s^n(\mathbf{w}_s^n, dz_s)\Big]\Big|_{\substack{l_s=L_s^n,\ \underline{z}_{s-1}^n=\underline{\mathbb{Z}}_{s-1}^n \\ \mathbf{w}_s^n=\mathbf{W}_s^n}} 1_{\{L_s^n\in\mathcal{A}_{s,c}^n\}}\Big] \\
= \; & E_{s-1}\Big[\int f_{s,c}^n\big(\underline{L}_s^n, \underline{\mathbb{Z}}_{s-1}^n, z_s\big)P_s^n(\mathbf{W}_s^n, dz_s)1_{\{L_s^n\in\mathcal{A}_{s,c}^n\}}\Big] \qquad (31)
\end{aligned}
$$

With (30), we obtain $\mathrm{E}_{s-1}\Big[f_{s,c}^n\prod_{s'=1}^s 1$

$\prod_{s'=1}^s 1_{\{L}\Big]$ $Due to (Lemma 2 gives a backwards formula for$ $\overline{\mathbf{E}}_n$ of (22). *Suppose that* $f\big(\underline{L}_S^n, \underline{\mathbb{Z}}_S^n\big)$ *is integrable. Then*

Theorem 3. **(a)** $\mathbf{E}_n = f_0^n$, *where* f_0^n *is given by (28) and the functions* f_s^n *for* $t\in\mathbb{T}$ *are recursively given by (27).*

(b) $\overline{\mathbf{E}}_n = E\big[f\big(\underline{L}_S^n, \underline{\mathbb{Y}}_S^n\big)\big] = \mathbf{E}_n + \rho_n$ *with*

$$
|\rho_n| \;\le\; E\Big[\big|f\big(\underline{L}_S^n, \underline{\mathbb{Z}}_S^n\big)\big|\sum_{s\in\mathbb{S}} 1_{\{L_s^n\notin\mathcal{A}_s^n\}}\Big].
$$

Remark 3. It was possible to formally refrain from introducing the variables V_t^n by incorporating V_t^n into L_t^n. However, we kept V_t^n because they have a different role from L_t^n in the causal inference.

B A backward approximation scheme and the backward propagation of errors

Theorem 3 gave an approximation to the expected value $E\big[f\big(\underline{L}_S^n, \underline{\mathbb{Y}}_S^n\big)\big]$ with a backwards formula for $\mathbf{E}_n$. The value of $\mathbf{E}_n$ is inductively computable by the formula (27) if we know the distribution $P_s^n(\mathbf{w}_s^n, dz_s)$ completely. However, it rarely has a closed form, and hence we need some approximation to $P_s^n(\mathbf{w}_s^n, dz_s)$ by a signed measure $\Psi_{s,\mathsf{p},\mathbf{w}_s^n}^n$ on $\mathbb{R}^{\mathsf{d}_s}$ depending on $\mathbf{w}_s^n$. Asymptotic expansion

method is promising in this situation. In fact, for batched bandits for example, in Section C, we will apply an expansion scheme under partial mixing (Yoshida [18]) to derive asymptotic expansion of $\mathbf{E}_n$. Before carrying out this plot, we shall prepare a framework of our method in this subsection.

Denote by $\underline{\mathcal{E}}_t(M,\gamma)$ the set of measurable functions $\mathsf{g} : \mathbb{R}^{\sum_{s=1}^t \mathsf{d}_s} \to \mathbb{R}$ such that

$$\big|\mathsf{g}(\underline{z}_t)\big| \quad \leq \quad M(1+|\underline{z}_t|^\gamma) \qquad \big(\underline{z}_t \in \mathbb{R}^{\sum_{s=1}^t \mathsf{d}_s}\big).$$

Suppose that

$$\big\{f^n_{S,c};\, \underline{l}_S \in \mathcal{L}\big\} \quad \subset \quad \underline{\mathcal{E}}_S(M_S,\gamma)$$

for some $M_S, \gamma > 0$. Denote by $\mathcal{E}_s(M,\gamma)$ the set of measurable functions $f : \big(\Pi_{s'=1}^t \mathcal{L}_{s'}\big) \times \mathbb{R}^{\sum_{s'=1}^t \mathsf{d}_{s'}} \to \mathbb{R}$ such that

$$\big|f(\underline{l}_t, \underline{z}_t)\big| \quad \leq \quad M(1+|\underline{z}_t|^\gamma) \qquad \big(\underline{l}_t \in \Pi_{s=1}^t \mathcal{L}_s,\ \underline{z}_t \in \mathbb{R}^{\sum_{s=1}^t \mathsf{d}_s}\big).$$

Some condition for integrability of the integrand is necessary to validate this formula at batch s. Moreover, we will need to restrict the class of functions f since the error bound of the asymptotic expansion depends on a measure-theoretic modulus of continuity of the function f. (This is clear if one considers approximation of the distribution function of a binomial distribution by an Edgeworth expansion. The error cannot be smaller than the order of $n^{-1/2}$ with any continuous distribution function for approximation.) Use of such a measure-theoretic modulus of continuity of the function f is inevitable and standard in the theory of asymptotic expansion. The reader is referred to Bhattacharya and Rao [4] for a construction of the theory of asymptotic expansion for independent variables.

For $(M,\gamma) \in \mathbb{R}^2_+$, denote by $\mathcal{E}_s(M,\gamma)$ the set of measurable functions $\mathsf{f} : \mathbb{R}^{\mathsf{d}_s} \to \mathbb{R}$ such that

$$\big|\mathsf{f}(z_s)\big| \quad \leq \quad M(1+|z_s|^\gamma) \qquad \big(z_s \in \mathbb{R}^{\mathsf{d}_s}\big).$$

Let $r_0 = r_0(\underline{z}_0) = 1$ and $r_s(\underline{z}_s) = \prod_{s'=1}^t (1+|z_{s'}|^\gamma)^{-1}$ for $s \in \mathbb{S}$. Given $\mathbf{M} = (M_s)_{s\in\mathbb{S}} \in \mathbb{R}^S_+$, denote by $\mathcal{D}(\mathbf{M},\gamma)$ a family of measurable functions $f : \mathcal{L} \times \mathbb{R}^{\mathsf{m}} \to \mathbb{R}$ satisfying the following condition:

$$\mathcal{M}_s(f) \quad \subset \quad \mathcal{E}_s(M_s,\gamma) \tag{32}$$

for every $s \in \mathbb{S}$, where

$$\mathcal{M}_s(f) \quad = \quad \Big\{ r_{s-1}(\underline{z}_{s-1})\big(|f^n_s| + |\widehat{f}^n_s|\big)\big(\underline{l}_s, \underline{z}_{s-1}, \cdot\big) 1_{\{l_s \in \mathcal{A}^n_s\}};\ \underline{l}_s, \underline{z}_{s-1}, n \Big\}.$$

Then, obviously, the functions $f^n_{s-1}(\underline{l}_{s-1}, \underline{z}_{s-1})$ and $\widehat{f}^n_{s-1}(\underline{l}_{s-1}, \underline{z}_{s-1})$ are inductively well defined if

$$\int |z_s|^\gamma \big\{P^n_s(\mathbf{w}^n_s, dz_s) + |\Psi^n_{s,\mathsf{p},\mathbf{w}^n_s}|(dz_s)\big\} 1_{\{l_s \in \mathcal{A}^n_s\}} \nu^n_{c_s}(dl_s, d\mathbf{w}^n_s) q\big((\underline{l}_{s-1}, \underline{z}_{s-1}), dc_s\big) < \infty \tag{33}$$

for every $\underline{l}_{s-1}, \underline{z}_{s-1}, s, n$. More precisely, under the integrability condition (33), if $\big\{f^n_{t,c}(\underline{l}_t, \cdot);\ \underline{l}_t \in \Pi_{s=1}^t \mathcal{L}_s\big\} \subset \underline{\mathcal{E}}_t(M_t,\gamma)$ for some positive constants M_t and γ, then $\big\{f^n_{t-1,c}(\underline{l}_{t-1}, \cdot);\ \underline{l}_{t-1} \in \Pi_{s=1}^{t-1} \mathcal{L}_s\big\} \subset \underline{\mathcal{E}}_{t-1}(M_{t-1},\gamma)$ for some positive constant M_{t-1}. The class $\mathcal{D}(\mathbf{M},\gamma)$ will be more precisely specified later by putting conditions that ensure a measure-theoretic continuity of functions.

For $f \in \mathcal{D}(M,\gamma)$ satisfying (32), the functions $f^n_{t-1,c}(\underline{l}_{t-1}, \underline{z}_{t-1})$ are inductively well defined if

$$\int |z_t|^\gamma P^n_s(\mathbf{w}^n_s, dz_t) 1_{\{l_t \in \mathcal{A}^n_{t,c}\}} \nu^n_{c_t}(dl_t, d\mathbf{w}^n_s) \quad < \quad \infty$$

for $t \in \mathbb{T}$.

then

Obviously, for positive constants M_t and γ, if

$$\big\{f^n_{t,c}(\underline{l}_t, \underline{z}_{t-1}, \cdot);\ \underline{z}_{t-1} \in \mathbb{R}^{\sum_{s=1}^{t-1} \mathsf{d}_s},\ \underline{l}_t \in \Pi_{s=1}^t \mathcal{L}_s\big\} \subset \mathcal{E}_t(M_t,\gamma)$$

and

and

$$\left\{ f^n_{t-1,c}(\underline{l}_{t-1}, \underline{z}_{t-2}, \cdot);\ \underline{l}_{t-1} \in \Pi_{s=1}^{t-1} \mathcal{L}_s \right\} \subset \mathcal{E}_{t-1}(M_{t-1}, \gamma)$$

for some positive constant M_{t-1}. The condition that $\left\{ f^n_{t,c}(\underline{l}_t, \cdot);\ \underline{l}_t \in \Pi_{s=1}^{t} \mathcal{L}_s \right\} \subset \underline{\mathcal{E}}_t(M_t, \gamma)$ is not restrictive. If this family of functions f behaves badly, then we can replace it by a tame one, by means of truncation to f, at (25) before running the reccursion.

$$\left\| E\big[f^n_{t,c}(\underline{z}_{t-1}, z_t) 1_{\left\{(l_t, z_t) \in B(c_{t+1})\right\}} P^n_s(\mathbf{w}^n_s, dz_t) \big] - \int f^n_{t,c}(\underline{z}_{t-1}, z_t) 1_{\left\{(l_t, z_t) \in B(c_{t+1})\right\}} \Psi_{n,p,\mathbf{w}^n_s}(dz_t) \right\|_{L^1(P)}$$

Define $\widetilde{f}^n_s(\underline{l}_s, \underline{z}_s)$ by $\widetilde{f}^n_S = f^n_S$ for f^n_S of (26) and

$$\begin{aligned} \widetilde{f}^n_{s-1}(\underline{l}_{s-1}, \underline{z}_{s-1}) \quad = \quad & \int_{\mathfrak{C}_s} \int_{\mathcal{L}_s \times \mathbb{R}^{\mathsf{d}_s}} f^n_s\big(\underline{l}_{s-1}, l_t, \underline{z}_{s-1}, z_s\big) \Psi^n_{s,\mathsf{p},\mathbf{w}^n_s}(dz_s) 1_{\{l_s \in \mathcal{A}^n_s\}} \nu^n_{c_s}(dl_s, d\mathbf{w}^n_s) \\ & \times q\big((\underline{l}_{s-1}, \underline{z}_{s-1}), dc_s\big) \end{aligned}$$

for $s \in \mathbb{S}$. We notice that the functions $\widetilde{f}^n_s(\underline{l}_s, \underline{z}_s)$ are well defined under the conditions (32) and (33).

Let

$$\varrho^n_s(f) \quad = \quad \sup_{\underline{l}_{s-1}, \underline{z}_{s-1}} \left| r_{s-1}(\underline{z}_{s-1}) f^n_{s-1}(\underline{l}_{s-1}, \underline{z}_{s-1}) - r_{s-1}(\underline{z}_{s-1}) \widetilde{f}^n_{s-1}(\underline{l}_{s-1}, \underline{z}_{s-1}) \right|$$

and

$$\widehat{\varrho}^n_s(f) \quad = \quad \sup_{\underline{l}_{s-1}, \underline{z}_{s-1}} \left| r_{s-1}(\underline{z}_{s-1}) f^n_{s-1}(\underline{l}_{s-1}, \underline{z}_{s-1}) - r_{s-1}(\underline{z}_{s-1}) \widehat{f}^n_{s-1}(\underline{l}_{s-1}, \underline{z}_{s-1}) \right|$$

for $s = 1, ..., S$. By definition, $\widehat{\varrho}^n_S(f) \ = \ \varrho^n_S(f)$.

We have

$$\begin{aligned} & r_{S-1}(\underline{z}_{S-1}) f^n_{S-1}(\underline{l}_{S-1}, \underline{z}_{S-1}) - r_{S-1}(\underline{z}_{S-1}) \widehat{f}^n_{S-1}(\underline{l}_{S-1}, \underline{z}_{S-1}) \\ \equiv \quad & r_{S-1}(\underline{z}_{S-1}) f^n_{S-1}(\underline{l}_{S-1}, \underline{z}_{S-1}) - r_{S-1}(\underline{z}_{S-1}) \widetilde{f}^n_{S-1}(\underline{l}_{S-1}, \underline{z}_{S-1}) \\ = \quad & \int r_{S-1}(\underline{z}_{S-1}) f\big(\underline{l}_{S-1}, l_S, \underline{z}_{S-1}, z_S\big) 1_{\{l_S \in \mathcal{A}^n_S\}} \\ & \quad \times \big\{ P^n_S(\mathbf{w}^n_S, dz_S) - \Psi^n_{S,\mathsf{p},\mathbf{w}^n_S}(dz_S) \big\} \nu^n_{c_S}(dl_S, d\mathbf{w}^n_S) q\big((\underline{l}_{s-1}, \underline{z}_{s-1}), dc_s\big) \end{aligned}$$

Then

$$
\begin{aligned}
\widehat{\varrho}_s^n(f) &= \sup_{\underline{l}_{s-1},\underline{z}_{s-1}} \left| r_{s-1}(\underline{z}_{s-1}) f_{s-1}^n(\underline{l}_{s-1},\underline{z}_{s-1}) - r_{s-1}(\underline{z}_{s-1}) \widehat{f}_{s-1}^n(\underline{l}_{s-1},\underline{z}_{s-1}) \right| \\
&\leq \sup_{\underline{l}_{s-1},\underline{z}_{s-1}} \left| r_{s-1}(\underline{z}_{s-1}) f_{s-1}^n(\underline{l}_{s-1},\underline{z}_{s-1}) - r_{s-1}(\underline{z}_{s-1}) \widetilde{f}_{s-1}^n(\underline{l}_{s-1},\underline{z}_{s-1}) \right| \\
&\quad + \sup_{\underline{l}_{s-1},\underline{z}_{s-1}} \left| r_{s-1}(\underline{z}_{s-1}) \widetilde{f}_{s-1}^n(\underline{l}_{s-1},\underline{z}_{s-1}) - r_{s-1}(\underline{z}_{s-1}) \widehat{f}_{s-1}^n(\underline{l}_{s-1},\underline{z}_{s-1}) \right| \\
&\leq \varrho_s^n(f) \\
&\quad + \sup_{\underline{l}_{s-1},\underline{z}_{s-1}} \left| \int \frac{r_{s-1}(\underline{z}_{s-1})}{r_s(\underline{z}_s)} \big\{ r_s(\underline{z}_s) f_s^n\big(\underline{l}_s,\underline{z}_s\big) - r_s(\underline{z}_s) \widehat{f}_s^n\big(\underline{l}_s,\underline{z}_s\big) \big\} \right. \\
&\qquad \left. \times \Psi_{s,\mathsf{p},\mathbf{w}_s^n}^n(dz_s) 1_{\{l_s \in \mathcal{A}_s^n\}} \nu_{c_s}^n(dl_s, d\mathbf{w}_s^n) q\big((\underline{l}_{s-1},\underline{z}_{s-1}), dc_s\big) \right| \\
&= \varrho_s^n(f) + \sup_{\underline{l}_{s-1},\underline{z}_{s-1}} \left| \int (1+|z_s|^\gamma) \big\{ r_s(\underline{z}_s) f_{s,c}^n\big(\underline{l}_s,\underline{z}_s\big) - r_s(\underline{z}_s) \widehat{f}_{s,c}^n\big(\underline{l}_s,\underline{z}_s\big) \big\} \right. \\
&\qquad \left. \times \Psi_{s,\mathsf{p},\mathbf{w}_s^n}^n(dz_s) 1_{\{l_s \in \mathcal{A}_s^n\}} \nu_{c_s}^n(dl_s, d\mathbf{w}_s^n) q\big((\underline{l}_{s-1},\underline{z}_{s-1}), dc_s\big) \right| \\
&\leq \varrho_s^n(f) \\
&\quad + \widehat{\varrho}_{s+1}^n(f) \sup_{\underline{l}_{s-1},\underline{z}_{s-1}} \int (1+|z_s|^\gamma) |\Psi_{s,\mathsf{p},\mathbf{w}_s^n}^n|(dz_s) 1_{\{l_s \in \mathcal{A}_s^n\}} \nu_{c_s}^n(dl_s, d\mathbf{w}_s^n) q\big((\underline{l}_{s-1},\underline{z}_{s-1}), dc_s\big) \\
&= \varrho_s^n(f) + \widehat{\varrho}_{s+1}^n(f) U_{s+1}^n,
\end{aligned}
$$

where

$$
U_{s+1}^n = \sup_{\underline{l}_{s-1},\underline{z}_{s-1}} \int (1+|z_s|^\gamma) |\Psi_{s,\mathsf{p},\mathbf{w}_s^n}^n|(dz_s) 1_{\{l_s \in \mathcal{A}_s^n\}} \nu_{c_s}^n(dl_s, d\mathbf{w}_s^n) q\big((\underline{l}_{s-1},\underline{z}_{s-1}), dc_s\big) \tag{34}
$$

for $s = 1, ..., S-1$. Thus,

$$
\begin{aligned}
\widehat{\varrho}_S^n(f) &= \varrho_S^n(f), \\
\widehat{\varrho}_{S-1}^n(f) &\leq \varrho_{S-1}^n(f) + \widehat{\varrho}_S^n(f) U_S^n, \\
\widehat{\varrho}_{S-2}^n(f) &\leq \varrho_{S-2}^n(f) + \widehat{\varrho}_{S-1}^n(f) U_{S-1}^n, \\
&\cdots \\
|f_{0,c}^n - \widehat{f}_{0,c}^n| = \widehat{\varrho}_1^n(f) &\leq \varrho_1^n(f) + \widehat{\varrho}_2^n(f) U_2^n.
\end{aligned}
$$

Therefore, the following result has been shown.

Theorem 4. *Suppose that the condition (33) is satisfied. Then*

$$
|f_{0,c}^n - \widehat{f}_{0,c}^n| \leq \sum_{s \in \mathbb{S}} \left(\varrho_s^n(f) \prod_{i=2}^s U_s^n \right) \tag{35}
$$

for all $f \in \mathcal{D}(\mathbf{M}, \gamma)$, $\mathbf{M} = (M_s)_{s \in \mathbb{S}}$, *where* $\prod_{i=2}^1 U_s^n = 1$.

B.1 Partial mixing and asymptotic expansion

Theorem 4 gave the estimate (35) of the error caused by approximation of $P_s^n(\mathbf{w}_s^n, dz_s)$ by $\Psi_{s,\mathsf{p},\mathbf{w}_s^n}^n(dz_s)$. It is possible to asses this error bound by means of the theory of asymptotic expansion under partial mixing. We begin with a generalization of the theory. [5]

[5] The letter T in this section denotes the total time or number of observations, e.g. n_s, not the number of batchs in the previous sections. We keep T for the reader's convenience since the proof in Section 5 deeply relies on the one in Yoshida [18], where T is used for the same role.

B.1.1 Asymptotic expansion for a partially mixing process

In this section, we recall the asymptotic expansion scheme for a partial mixing process given by Yoshida [18]). We will present it in a simplified form of the theory though the paper was written in more generality. Moreover, this section includes a slight generalization of the original configuration, for this paper's use. For the reader's convenience for consulting the paper, we will basically keep the notation in Yoshida [18]). For example, the symbol "$\mathcal{C}$" is used here to denote the conditioning σ-field for partial mixing, not the set of the strategies in the previous sections.

On a probability space $(\Omega, \mathcal{F}, P)$, we consider a family of sub σ-fields $(\mathcal{F}_I)_{I\subset\mathbb{R}_+}$ such that $\mathcal{B}_I \subset \mathcal{G}_J$ for $I \subset J \subset \mathbb{R}_+$. Suppose that a sub σ-field $\mathcal{C}$ of $\mathcal{F}$ and $\mathcal{C}$-measurable random numbers $\alpha(s,t|\mathcal{C})$ $(s,t \in \mathbb{R}_+, s \le t)$ satisfy

$$1 \;\ge\; \alpha(s,t|\mathcal{C}) \;\ge\; \sup\left\{\left|P_{\mathcal{C}}[B_1\cap B_2] - P_{\mathcal{C}}[B_1]P_{\mathcal{C}}[B_2]\right|;\; B_1 \in \mathcal{B}_{[0,s]\vee\mathcal{C}}, B_2 \in \mathcal{B}_{[t,\infty)\vee\mathcal{C}}\right\}.$$

Here $P_{\mathcal{C}}$ denotes a version of conditional probability given $\mathcal{C}$. The random field $\alpha(s,t|\mathcal{C})$ is called a $\mathcal{D}$-conditional α-mixing coefficient. We suppose that $\mathcal{C}$-measurable random numbers $\alpha(h|\mathcal{C})$ $(h \in \mathbb{R}_+)$ satisfy

$$1 \;\ge\; \alpha(h|\mathcal{C}) \;\ge\; \sup\left\{\alpha(t,t+h'|\mathcal{C});\; h' \ge h, t \in \mathbb{R}_+\right\}$$

for $h \in \mathbb{R}_+$.

We will consider a d-dimensional process $Z = (Z_t)_{t\in\mathbb{R}_+}$ that Z_0 is $\mathcal{B}_{[0]} \vee \mathcal{C}$-measurable and that $Z_t^s := Z_t - Z_s$ is $\mathcal{B}_{[s,t]} \vee \mathcal{C}$-measurable for every $s,t \in \mathbb{R}_+$, $s \le t$. The increment of Z on the interval $I \subset \mathbb{R}_+$ is denoted by Z_I, i.e., $Z_I = Z_{\sup I} - Z_{\inf I}$.

Given a positive integer number $\mathsf{p} \ge 3$, the $(\mathsf{p}-2)$-th asymptotic expansion formula for the distribution $T^{-1/2}Z_T$ is given by the $\mathcal{C}$-conditional formula $\Psi_{T,\mathsf{p},\mathcal{C}}$. The formula $\Psi_{T,\mathsf{p},\mathcal{C}}$ is a $\mathcal{C}$-measurable random signed measure specified as follows. The $\mathcal{C}$-conditional cummulant functions $\chi_{T,r,\mathcal{C}}(u)$ of $T^{-1/2}Z_T$ are defined by

$$\chi_{T,r,\mathcal{C}}(u) \;=\; \left(\frac{\partial}{\partial\epsilon}\right)^r \log E_{\mathcal{C}}\left[\exp\left(\mathtt{i}\epsilon u\cdot T^{-1/2}Z_T\right)\right]\Big|_{\epsilon=0} \tag{36}$$

for $u \in \mathbb{R}^{\mathsf{d}}$, where $E_{\mathcal{C}}$ denotes the conditional expectation given $\mathcal{C}$. Next, the $\mathcal{C}$-random functions $P_{T,r,\mathcal{C}}(u)$ of $u \in \mathbb{R}^{\mathsf{d}}$ are defined by the formal expansion

$$\exp\left(\sum_{r=2}^{\infty}\epsilon^{r-2}(r!)^{-1}\chi_{T,r,\mathcal{C}}(u)\right) \;=\; \exp\left(2^{-1}\chi_{T,2,\mathcal{C}}(u)\right)\sum_{r=0}^{\infty}\epsilon^r T^{-r/2}P_{T,r,\mathcal{C}}\left(u;\chi_{T,\cdot,\mathcal{C}}\right), \tag{37}$$

$P_{T,0,\mathcal{C}}\left(u;\chi_{T,\cdot,\mathcal{C}}\right) = 1$, and the function $\widehat{\Psi}_{T,\mathsf{p},\mathcal{C}}(u)$ of $u \in \mathbb{R}^{\mathsf{d}}$ is defined as the partial sum

$$\widehat{\Psi}_{T,\mathsf{p},\mathcal{C}}(u) \;=\; \exp\left(2^{-1}\chi_{T,2,\mathcal{C}}(u)\right)\sum_{r=0}^{\mathsf{p}-2}T^{-r/2}P_{T,r,\mathcal{C}}\left(u;\chi_{T,\cdot,\mathcal{C}}\right). \tag{38}$$

Finally the random signed measure $\Psi_{T,\mathsf{p},\mathcal{C}}$ is defined as the Fourier inversion of $\widehat{\Psi}_{T,\mathsf{p},\mathcal{C}}$:

$$\Psi_{T,\mathsf{p},\mathcal{C}} \;=\; \mathcal{F}^{-1}\left[\widehat{\Psi}_{T,\mathsf{p},\mathcal{C}}\right]. \tag{39}$$

We set $\Psi_{T,\mathsf{p},\mathcal{C}} = \delta_0$ if the quadratic form $\mathrm{Var}_{\mathcal{C}}[T^{-1/2}Z_T][u^{\otimes 2}] = -\chi_{T,2,\mathcal{C}}(u)$ degenerates. [It should be remarked that, in the proof of Theorem 5, the signed measure $\Psi_{T,\mathsf{p},\mathcal{C}}$ defined by the Fourier inversion (39) is only used on the event $\{s_T I_{\mathsf{d}} \le \mathrm{Var}_{\mathcal{C}}[T^{-1/2}Z_T] \le u_T I_{\mathsf{d}}\}$. So this proviso is not essential in the story.]

The error of the formula is evaluated by

$$\Delta_T(\mathsf{f}) \;=\; \left\|E_{\mathcal{C}}\left[\mathsf{f}\left(T^{-1/2}Z_T\right)\right] - \Psi_{T,\mathsf{p},\mathcal{C}}[\mathsf{f}]\right\|_1 \tag{40}$$

In what follows, it is assumed that

$$\limsup_{T\to\infty}\left\|\left|\Psi_{T,\mathsf{p},\mathcal{C}}\right|[1+|\cdot|^{\mathsf{p}_0}]\right\|_q \;<\; \infty \tag{41}$$

for some $q > 1$, where $\mathsf{p}_0 = 2\lfloor \mathsf{p}/2 \rfloor$. In (40), we consider $\mathcal{C}$-measurable random fields f, generalizing the results in Yoshida [18]. This extension is necessary for the application to the causal inference treated in this paper. Introduce a $\mathcal{C}$-measurable truncation functional $\Phi_T : \Omega \to \{0,1\}$, and define a truncated modulus $\omega_2^{\Phi}(\mathsf{f};\epsilon,\nu)$ of continuity of the random field f by

$$\omega_2^{\Phi}(\mathsf{f};\epsilon,\nu) \quad = \quad \left\{ E\left[\Phi_T \int_{\mathbb{R}^{\mathsf{d}}} \omega_{\mathsf{f}}(x;\epsilon)^2 \nu(dx) \right] \right\}^{1/2}, \tag{42}$$

where $\omega_{\mathsf{f}}(x;\epsilon) = \sup\left\{ |\mathsf{f}(x+h) - \mathsf{f}(x)|;\ |h| \leq \epsilon \right\}$ and $\mathcal{C}$-measurability of this modulus of continuity is implicitly assumed for f when we consider $\omega_{\mathsf{f}}(x;\epsilon)$. We assume that

$$P\big[\Phi_T = 0\big] \quad = \quad O(T^{-L}) \tag{43}$$

as $T \to \infty$ for every $L > 0$. Denote by $\check{\mathcal{E}}(M,\gamma)$ the set of functions $\mathsf{f} : \Omega \times \mathbb{R}^{\mathsf{d}} \to \mathbb{R}$ that are $\mathcal{C} \times \mathbb{B}_{\mathsf{d}}$-measurable and satisfy

$$|\mathsf{f}(\omega,x)| \quad \leq \quad M\big(1 + |x|^{\gamma}\big) \quad (\omega \in \Omega,\ x \in \mathbb{R}^{\mathsf{d}}).$$

For the sequence $\Phi = (\Phi_T)_{T \in \mathbb{R}_+}$ of $\mathcal{C}$-measurable truncation functionals, we denote by $\mathcal{E}^{\Phi}(M,\gamma)$ koko

$$1_{\{\Phi_T(\omega)=0\}}\omega.....koko$$

We will work with the following conditions.

[PM1] There exists a positive constant a such that $\|\alpha(h|\mathcal{C})\|_1 \leq a^{-1}\exp(-ah)$ for all $h > 0$.

[PM2] There exists a positive number h_0 such that for every $L \in \mathbb{N}$,

$$E\big[\big(E_{\mathcal{C}}[|Z_0|]|\big)^L\big] + \sup_{\substack{t,h:t\in\mathbb{R}_+ \\ 0\leq h\leq h_0}} E\left[\big(E_{\mathcal{C}}\big[|Z_{t+h}^t|^{\mathsf{p}+1}\big]\big)^L\right] \quad < \quad \infty$$

and $E_{\mathcal{C}}[Z_t] = 0$ a.s. for all $t \in \mathbb{R}_+$.

To prove an error bound in asymptotic expansion, we need some regularity of the distribution of the functional in question. The so-called Cramér condition serves in the classical setting of independent variables to validate the Fourier inversion of the characteristic function of the sum. By the independency, the characteristic function is factorized into that of each increment and the Cramér condition provides a decay of each factor. For mixing processes, if the Markovian property is available, a similar factorization is feasible with the aid of the conditional characteristic function of each increment, and the Cramér condition is replaced by a conditional version of it. For functionals in stochastic analysis, a local non-degeneracy of the Malliavin covariance matrix of the increments of the functional undertakes this role ([18]).

The notion of reduction intervals is necessary. For a collection of intervals $(I_\ell)_{\ell\in\{1,\ldots,n'(T)\},\,T\in\mathbb{N}}$ is called (dense) reduction intervals associated with $\mathcal{C}$ and $\big(\widehat{\mathcal{C}}(\ell)\big)_{\ell\in\{1,\ldots,n'(T)\},\,T\in\mathbb{N}}$ if the following conditions are satisfied.

[(i)]$\widehat{\mathcal{C}}(\ell)$ ($\ell \in \{1,...,n'(T)\}$, $T \in \mathbb{N}$) are sub σ-fields of $\mathcal{F}$ with $\widehat{\mathcal{C}}(\ell) \supset \mathcal{C}$. $I_\ell = [u_\ell, v_\ell]$ satisfies $v_\ell \leq u_{\ell+1}$ for $\ell \in \{1,...,n'(T)-1\}$, $0 < \inf_{T,\ell}\big(v_\ell - u_\ell\big) \leq \sup_\ell\big(v_\ell - u_\ell\big) < \infty$, and $\liminf_{T\to\infty}\big(n'(T)/T\big) > 0$. For any subsequence $\ell_1,...,\ell_{n''(T)}$ of $1,...,n'(T)$, any bounded $\mathcal{B}_{[0,T]\setminus\cup_{\ell\in\{\ell_1,\ldots,\ell_{n''(T)}\}}I_\ell} \vee \mathcal{C}$-measurable random variable A and any bounded $\mathcal{B}_{I_{\ell''}} \vee \mathcal{C}^s$-measurable random variable $\mathsf{A}_{\ell''}$, $\ell'' \in \{\ell_1,...,\ell_{n''(T)}\}$, it holds that

$$E_{\mathcal{C}}\big[\mathsf{A}\mathsf{A}_{\ell_1}\cdots\mathsf{A}_{\ell_{n''(T)}}\big] \quad = \quad E_{\mathcal{C}}\left[\mathsf{A}E_{\widehat{\mathcal{C}}(\ell_1)}[\mathsf{A}_{\ell_1}]\cdots E_{\widehat{\mathcal{C}}(\ell_{n''(T)})}[\mathsf{A}_{\ell_{n''(T)}}]\right]. \tag{44}$$

In the above conditions, I_ℓ and $\widehat{\mathcal{C}}(\ell)$ may depend on T. The condition (44) is a conditional type of Markovian property. For example, if the underlying process of the system is Markovian, then we can take the σ-field $\widehat{\mathcal{C}}^s_{T,\ell}$ consisting of the information of the process at the boundary of $I^s_{T,\ell}$.

To the present case, the classical Cramér condition is extended in a somewhat abstract way by the following condition.

[PM3] For every $L > 0$, there exist truncation functionals $\psi_j : (\Omega, \mathcal{F}) \to ([0,1], \mathbb{B}([0,1]))$ and constants $a', a \in (0,1)$ and $B > 0$ such that $4a' < (a-1)^2$ and

(i) $P\left[n'(T)^{-1} \sum_j E_{\mathcal{C}}\left[\sup_{u:|u|\geq B} \left|E_{\widehat{\mathcal{C}}(j)}\left[e^{\mathrm{i}u\cdot Z_{I(j)}}\psi_j\right]\right|\right] \geq a'\right] = o(T^{-L})$ as $T \to \infty$.

(ii) $P\left[n'(T)^{-1} \sum_j P_{\mathcal{C}}[1-\psi_j] \geq a\right] = o(T^{-L})$ as $T \to \infty$.

[PM3] For every $L > 0$, there exist truncation functionals $\psi_j : (\Omega, \mathcal{F}) \to ([0,1], \mathbb{B}([0,1]))$, and there exist positive constants η_1, η_2, η_3 and B such that $\eta_1 + \eta_2 < 1$, $\eta_3 < 1$ and that

$$P\left[\sum_j P_{\mathcal{C}}\left[\widetilde{\Phi}(j) \leq 1-\eta_2\right] > \eta_3 n'(T)\right] \quad = \quad o(T^{-L})$$

as $T \to \infty$, where

$$\widetilde{\Phi}(j) \quad = \quad E_{\widehat{\mathcal{C}}(j)}\left[\psi_j\right] \cdot 1_{\left\{\sup_{u:|u|\geq B}\left|E_{\widehat{\mathcal{C}}(j)}\left[\psi_j \exp(\mathrm{i}u\cdot Z_{I(j)})\right]\right| < \eta_1\right\}}.$$

Condition $[PM3]$ is slihtly stronger than the usual condition in independent cases. As a matter of fact, it is possible to weaken this condition; see Remark 15 of Yoshida [18], p.609. Condition $[PM3]$ is Condition $[A3^\natural]$ mentioned by Remark 15 of Yoshida [18], p.609.

The following result is a generalization of Theorem 1 of Yoshida [18]. We will treat the conditionally exponentially mixing case to give priority to simplicity though slower mixing cases discussed in the above paper could be treated. Write $I^\varrho = \{t \in \mathbb{R}_+;\ \mathrm{dist}(t, I) \leq \varrho\}$.

Theorem 5. *Let $\left(I(j)\right)_{j=1,\ldots,n'(T)}$ be dense reduction intervals with $\left(\widehat{\mathcal{C}}(j)\right)_{j=1,\ldots,n'(T)}$. Suppose that there exists a positive constant ϱ such that $\widehat{\mathcal{C}}(j) \subset \mathcal{B}_{I(j)^\varrho} \vee \mathcal{C}$. Suppose that Conditions $[PM1]$, $[PM2]$ and $[PM3]$ are satisfied. Let $K, M \in (0,\infty)$, $\theta \in (0,1)$, and let s_T, u_T be any sequences of positive numbers with $\liminf_{T\to\infty} u_T/T^{c''} > 0$ for some positive constant c''. Then there exist positive constants c', M^* and δ^* such that*

$$\begin{aligned}\sup_{\mathsf{f}\in\check{\mathcal{E}}(M,\mathsf{p}_0)} \Delta_T(\mathsf{f}) \quad \leq \quad & M^*\Bigg\{\left(P\left[Var_{\mathcal{C}}\left[T^{-1/2}Z_T\right] < s_T I_{\mathsf{d}}\right]\right)^\theta \\ & \qquad + u_T^{\gamma(1)} s_T^{-\gamma(2)} \sup_{\mathsf{f}\in\check{\mathcal{E}}(M,\gamma)} \omega_2^\Phi\left(\mathsf{f}; T^{-K}, \phi(x;0,u_T I_{\mathsf{d}})\right)\Bigg\} \\ & + o(T^{-(\mathsf{p}-2+\delta^*)/2})\end{aligned} \tag{45}$$

as $T \to \infty$, if $s_T \geq T^{-c'}$ for large T, where $\gamma(1) = [3(\mathsf{p}-2)+\mathsf{d}]/2$ and $\gamma(2) = 3(\mathsf{p}-2)\mathsf{d} + \mathsf{d}/2$.

Remark 4. In application of Theorem 5 to a not so bad class of f such as $\omega_2^\Phi\left(\mathsf{f}; T^{-K}, \phi(x;0,u_T I_{\mathsf{d}})\right) = O(T^{-\epsilon K})$ for some positive number ϵ, we can set a number for $K > \gamma(1)c'' + 1 + (\mathsf{p}-2+1)/2$, $u_T = T^{c''}$, and $s_T = T^{-c'_*}$ with a positive constant $c'_* < c'$ such that $c'_*\gamma(2) < 1$, and expect that the term $(P[....])^\theta$ on the right-hand side of (45) becomes as small as we like.

To prove Theorem 5, we follow the proof of Yoshida [18] but some technical modifications are necessary, as sketched in Section B.1.3.

B.1.2 Expansion of a transformed variable

Suppose that G_T is a $\mathsf{d} \times \mathsf{d}$ $\mathcal{C}$-measurable random matrix for $T \in \mathbb{R}_+$. Let

$$Z_T^* \ = \ T^{-1/2} G_T Z_T. \tag{46}$$

Suppose that G_T is invertible when $\Phi_T = 1$, and that

$$\sup_{T\in\mathbb{R}_+,\omega\in\Omega} \left(\Phi_T|G_T| + \Phi_T|G_T^{-1}|\right) \quad < \quad \infty. \tag{47}$$

We will use a $\mathcal{C}$-measurable random signed measure $\Psi^*_{T,\mathsf{p},\mathcal{C}}$ to approximate the distribution of Z^*_T. It is defined in a similar fashon in parallel to the procedure from (36) to (39). More precisely, let

$$\chi^*_{T,r,\mathcal{C}}(v) \quad = \quad \left(\frac{\partial}{\partial\epsilon}\right)^r \log E_{\mathcal{C}}\left[\exp\left(\mathtt{i}\epsilon v\cdot Z^*_T\right)\right]\Big|_{\epsilon=0} \tag{48}$$

for $v\in\mathbb{R}^{\mathsf{d}}$. Then, the function $\widehat{\Psi}^*_{T,\mathsf{p},\mathcal{C}}(v)$ of $v\in\mathbb{R}^{\mathsf{d}}$ is defined as the sum

$$\widehat{\Psi}^*_{T,\mathsf{p},\mathcal{C}}(v) \quad = \quad \exp\left(2^{-1}\chi^*_{T,2,\mathcal{C}}(v)\right)\sum_{r=0}^{\mathsf{p}-2} T^{-r/2}P_{T,r,\mathcal{C}}\left(v;\chi_{T,\cdot,\mathcal{C}}\right), \tag{49}$$

and the random signed measure $\Psi^*_{T,\mathsf{p},\mathcal{C}}$ is defined as the Fourier inversion of $\widehat{\Psi}^*_{T,\mathsf{p},\mathcal{C}}$:

$$\Psi^*_{T,\mathsf{p},\mathcal{C}} \quad = \quad \mathcal{F}^{-1}\left[\widehat{\Psi}^*_{T,\mathsf{p},\mathcal{C}}\right]. \tag{50}$$

Instead of (40), the error of approximation to Z^*_T by $\Psi^*_{T,\mathsf{p},\mathcal{C}}$ is measured by

$$\Delta^*_T(\mathsf{f}) \quad = \quad \left\|E_{\mathcal{C}}\left[\mathsf{f}\left(Z^*_T\right)\right]-\Psi^*_{T,\mathsf{p},\mathcal{C}}[\mathsf{f}]\right\|_1. \tag{51}$$

Suppose that (41) is satisfied.

Theorem 6. *Suppose that the same conditions as Theorem 5 are satisfied. there exist positive constants c', M^* and δ^* such that the inequality (45) holds true but the constant M^* is different from that of Theorem 5. Then there exist positive constants c', M^{**} and δ^* such that*

$$\begin{aligned}\sup_{\mathsf{f}\in\check{\mathcal{E}}(M,\mathsf{p}_0)}\Delta^*_T(\mathsf{f}) \quad \leq \quad & M^{**}\Bigg\{\left(P\left[\mathit{Var}_{\mathcal{C}}\left[T^{-1/2}Z_T\right]<s_TI_{\mathsf{d}}\right]\right)^{\theta} \\ & \qquad +u_T^{\gamma(1)}s_T^{-\gamma(2)}\sup_{\mathsf{f}\in\check{\mathcal{E}}(M,\gamma)}\omega_2^{\Phi}\left(\mathsf{f};T^{-K},\phi(x;0,u_TI_{\mathsf{d}})\right)\Bigg\} \\ & +o(T^{-(\mathsf{p}-2+\delta^*)/2})\end{aligned} \tag{52}$$

as $T\to\infty$, if $s_T\geq T^{-c'}$ for large T.

Proof. For $\mathsf{f}\in\check{\mathcal{E}}(M,\gamma)$, let

$$\mathsf{f}^*(z) \quad = \quad \Phi_T\mathsf{f}(G_Tz).$$

Then $\mathsf{f}^*\in\check{\mathcal{E}}(M',\gamma)$ for some positive constant M' by the assumption (47). We apply Theorem 5 to the $\mathcal{C}$-measurable random function f^* instead of f. We notice that the term $\omega_2^{\Phi}\left(\mathsf{f}^*;T^{-K},\phi(x;0,u_TI_{\mathsf{d}})\right)$ can be estimated by $\omega_2^{\Phi}\left(\mathsf{f};T^{-K},\phi(x;0,u_TI_{\mathsf{d}})\right)$ multiplied by a constant associated with the bound (47). □

From (36) and (48),

$$\chi^*_{T,r,\mathcal{C}}(v) \quad = \quad \chi_{T,r,\mathcal{C}}({G_T}^{\star}v). \tag{53}$$

By the definition of $P_{T,r,\mathcal{C}}$,

$$\exp\left(\sum_{r=2}^{\infty}\epsilon^{r-2}(r!)^{-1}\chi^*_{T,r,\mathcal{C}}(v)\right) \quad = \quad \exp\left(2^{-1}\chi^*_{T,2,\mathcal{C}}(v)\right)\sum_{r=0}^{\infty}\epsilon^rT^{-r/2}P_{T,r,\mathcal{C}}\left(v;\chi^*_{T,\cdot,\mathcal{C}}\right),$$

and we obtain

$$\exp\left(\sum_{r=2}^{\infty}\epsilon^{r-2}(r!)^{-1}\chi_{T,r,\mathcal{C}}(u)\right) \quad = \quad \exp\left(2^{-1}\chi_{T,2,\mathcal{C}}(u)\right)\sum_{r=0}^{\infty}\epsilon^rT^{-r/2}P_{T,r,\mathcal{C}}\left((G^{\star}_T)^{-1}u;\chi^*_{T,\cdot,\mathcal{C}}\right), \tag{54}$$

by plugging $(G^{\star}_T)^{-1}u$ into v and by using (53). Comparing (38) and (54), we see

$$P_{T,r,\mathcal{C}}\left(u;\chi_{T,\cdot,\mathcal{C}}\right) \quad = \quad P_{T,r,\mathcal{C}}\left((G^{\star}_T)^{-1}u;\chi^*_{T,\cdot,\mathcal{C}}\right),$$

or equivalently,

$$P_{T,r,\mathcal{C}}\big(G_T^\star v;\chi_{T,\cdot,\mathcal{C}}\big) \quad = \quad P_{T,r,\mathcal{C}}\big(v;\chi_{T,\cdot,\mathcal{C}}^*\big). \tag{55}$$

On the event

$$\begin{aligned} & \big\{\mathrm{Var}_{\mathcal{C}}[T^{-1/2}Z_T][u^{\otimes 2}] \text{ is non-degenerate }\big\}\cap\{\Phi_T=1\} \\ = \quad & \big\{\mathrm{Var}_{\mathcal{C}}[Z_T^*][u^{\otimes 2}] \text{ is non-degenerate }\big\}\cap\{\Phi_T=1\}, \end{aligned}$$

we have

$$\begin{aligned} \Psi_{T,\mathsf{p},\mathcal{C}}[\mathsf{f}^*] &= \int \Phi_T f(G_T z)\Psi_{T,\mathsf{p},\mathcal{C}}\mathsf{f}(dz) \\ &= \int \Phi_T f(G_T z)\frac{d\mathcal{F}^{-1}\big[\widehat{\Psi}_{T,\mathsf{p},\mathcal{C}}\big]}{dz}(z)dz \\ &= \int \Phi_T f(G_T z)\sum_{r=0}^{\mathsf{p}-2}T^{-r/2}P_{T,r,\mathcal{C}}\big(-\mathtt{i}\partial_z;\chi_{T,\cdot,\mathcal{C}}\big)\phi\big(z;0,\mathrm{Var}_{\mathcal{C}}[T^{-1/2}Z_T]\big)dz \\ &= \int \Phi_T f(G_T z)\sum_{r=0}^{\mathsf{p}-2}T^{-r/2}P_{T,r,\mathcal{C}}\big(G_T^\star(-\mathtt{i}\partial_y);\chi_{T,\cdot,\mathcal{C}}\big)\phi\big(G_T^{-1}y;0,\mathrm{Var}_{\mathcal{C}}[T^{-1/2}Z_T]\big)dz \end{aligned}$$

for $y=G_T z$, since

$$\frac{\partial}{\partial z} \quad = \quad G_T^\star\frac{\partial}{\partial y}.$$

Moreover,

$$\Psi_{T,\mathsf{p},\mathcal{C}}[\mathsf{f}^*] \quad = \quad \int \Phi_T f(G_T z)\sum_{r=0}^{\mathsf{p}-2}T^{-r/2}P_{T,r,\mathcal{C}}\big((-\mathtt{i}\partial_y);\chi_{T,\cdot,\mathcal{C}}^*\big)\phi\big(y;0,\mathrm{Var}_{\mathcal{C}}[Z_T^*]\big)|\det G_T|dz \quad (y=G_T z)$$

if (55) is used to rewrite the random symbols $P_{T,r,\mathcal{C}}\big(G_T^\star(-\mathtt{i}\partial_y);\chi_{T,\cdot,\mathcal{C}}\big)$. Therefore,

$$\begin{aligned} \Psi_{T,\mathsf{p},\mathcal{C}}[\mathsf{f}^*] &= \int \Phi_T f(y)\sum_{r=0}^{\mathsf{p}-2}T^{-r/2}P_{T,r,\mathcal{C}}\big((-\mathtt{i}\partial_y);\chi_{T,\cdot,\mathcal{C}}^*\big)\phi\big(y;0,\mathrm{Var}_{\mathcal{C}}[Z_T^*]\big)dy \\ & \qquad\qquad\qquad\qquad\qquad\qquad \text{(by change of variables from } z \text{ to } y=G_T z) \\ &= \Phi_T\Psi_{T,\mathsf{p},\mathcal{C}}^*[f] \end{aligned}$$

Now we apply Theorem 5 to f^* to prove the desired inequality. We remark that the contribution of $(1-\Phi_T)\mathsf{f}(Z_T^*)$ to the expectation is of $o(T^{-(\mathsf{p}-2+\delta^*)/2})$. □

Remark 5. Theorems 5 and 6 have been stated for the continuous parameter T. By simple embedding of a process $(V_n)_{n\in\mathbb{Z}_+}$ into a continuous time process as $(V_{\lfloor T\rfloor})_{T\in\mathbb{R}_+}$, we can apply these theorems to discrete time processes. It is also possible to reconstruct the theorems for a discrete time. Then the proof is unchaged.

B.1.3 A sketch of the proof of Theorem 5

We will only sketch a route to the result in Theorem 5 since the proof is almost the same as Section 6 of Yoshida [18], pp. 585-621, except for the treatment of $\mathcal{C}$-measurable functions f differently from it.

Fix some positive constants δ and $\widetilde{\delta}$ satisfying $\delta<\widetilde{\delta}$. We consider a sequence $\big(t_{T,j}\big)_{j=1,\ldots,N_T}$ for every $T\in\mathbb{R}_+$ such that $0=t_{T,0}<t_{T,1}<\cdots<t_{T,N_T}=T$ and $\delta<t_{T,j}-t_{T,j-1}<\widetilde{\delta}$ $(j=1,\ldots,N_T-1)$ and $t_{T,N_T}-t_{T,N_T-1}<\widetilde{\delta}$. Let $\widetilde{I}_j=[t_{T,j-1},t_{T,j}]$ $(j=1,\ldots,N_T)$. By definition, $Z_{\widetilde{I}_j}=Z_{t_{T,j}}-Z_{t_{T,j-1}}$ for $j=1,\ldots,N_T$. Let $Z_{\widetilde{I}_0}=Z_0$.

Let $\varphi : \mathbb{R}^{\mathsf{d}} \to [0,1]$ be a measurable function such that $\varphi(x) = 1$ if $|x| \leq 1/2$, and $\varphi(x) = 0$ if $|x| \geq 1$. The function φ_T is defined by $\varphi_T(x) = x\varphi(x/(2T^{\beta}))$, where $\beta \in (0, 1/2)$ is a constant; see pp.585-586 of Yoshida [18]. Let $Z'_T = \sum_{j=0}^{N_T} \varphi_T(Z_{\widetilde{I}_j})$ (p.601) and

$$e_T \quad = \quad T^{-1/2}\Big\{E_{\mathcal{C}}[\varphi_T(Z_0)] + \sum_{j=1}^{N_T} E_{\mathcal{C}}\big[\varphi_T\big(Z_{\widetilde{I}_j}\big)\big]\Big\}$$

as was done on p.596. Let $\widetilde{Z}_{T,j} = \varphi_T(Z_{\widetilde{I}_j}) - E_{\mathcal{C}}\big[\varphi_T(Z_{\widetilde{I}_j})\big]$ for $j = 0, 1, ..., N_T$. Furthermore, let $\widetilde{Z}_T = \sum_{j=1}^{N_T} \widetilde{Z}_{T,j}$ and $S^*_T = T^{-1/2}\widetilde{Z}_T$. Set $H_T(u, \mathcal{C}) = E_{\mathcal{C}}\big[e^{iu\cdot S^*_T}\big]$ for $u \in \mathbb{R}^{\mathsf{d}}$. As defined on p. 565 of [18], $\omega(\mathsf{f}; \epsilon, \nu) = \int \omega_{\mathsf{f}}(x; \epsilon)d\nu$ for a Borel measure ν on $\mathbb{R}^{\mathsf{d}}$.

On p.599 of Yoshida [18], the truncation functional ψ_T is defined as the indicator function fo the set

$$\Omega_0(T) \cap \big\{s_T I_{\mathsf{d}} \leq \mathrm{Var}_{\mathcal{C}}[T^{-1/2}Z_T] \leq u_T I_{\mathsf{d}}\big\},$$

where the event $\Omega_0(T)$ is specified on p.586. In place of Inequality (34) (p.605) in the proof of its Lemma 11 (p.601) of Yoshida [18], we now estimate

$$\Big\|\psi_T\Big(E_{\mathcal{C}}\big[\mathsf{f}(T^{-1/2}Z'_T)\big] - \Psi_{T,p,\mathcal{C}}[\mathsf{f}]\Big)\Big\|_1.$$

as follows. First,

$$\begin{aligned}
& \Big\|(1-\Phi_T)\psi_T\Big(E_{\mathcal{C}}\big[\mathsf{f}(T^{-1/2}Z'_T)\big] - \Psi_{T,p,\mathcal{C}}[\mathsf{f}]\Big)\Big\|_1 \\
\lesssim \quad & \big\|1-\Phi_T\big\|_1 T^{\gamma\beta} + \big\|1-\Phi_T\big\|_{q'}\big\|\Psi_{T,\mathsf{p},\mathcal{C}}[1+|\cdot|^{\mathsf{p}_0}]\big\|_q \\
= \quad & O(T^{-L})
\end{aligned}$$

as $T \to \infty$ for every $L > 0$, due to (41) and (43), where q is an index appearing in (41) and $q' = q/(q-1)$. Next, based on Sweeting's smoothing inequality applied (inside of the L^1-norm below) to the $\mathcal{C}$-measurable function f and the conditional probability measures $P_{\mathcal{C}}^{T^{-1/2}Z'_T}$ and $\Psi_{T,\mathsf{p},\mathcal{C}}$, we obatin an inequality corresponding to (34) of Yoshida [18]:

$$\begin{aligned}
& \Big\|\Phi_T\psi_T\Big(E_{\mathcal{C}}\big[\mathsf{f}(T^{-1/2}Z'_T)\big] - \Psi_{T,p,\mathcal{C}}[\mathsf{f}]\Big)\Big\|_1 \\
\leq \quad & CM\Big\{\Big\|\Phi_T\psi_T\int_{\mathbb{R}^{\mathsf{d}}} h(|x|)\big|\mathcal{K}_{T^{-K_1}} * \big(P_{\mathcal{C}}^{T^{-1/2}Z'_T} - \Psi_{T,\mathsf{p},\mathcal{C}}\big)\big|(dx)\Big\|_1 \\
& \quad + \Big\|\big(P_{\mathcal{C}}^{T^{-1/2}Z'_T} + |\Psi_{T,\mathsf{p},\mathcal{C}}|\big)[h(|\cdot|)]\Big\|_1 \int_{x:|x|\geq aT^{\zeta}} h(|x|)\mathcal{K}(dx) \\
& \quad + \Big\|\big(P_{\mathcal{C}}^{T^{-1/2}Z'_T} + |\Psi_{T,\mathsf{p},\mathcal{C}}|\big)[h(|\cdot|)]\Big\|_1 \delta(\alpha)^{\lfloor a^{-1/2}T^{\zeta}\rfloor}\Big\} \\
& + C'\Big\|\psi_T\Phi_T\int_{\mathbb{R}^{\mathsf{d}}}\sup_{x:|x|\leq aT^{-(K+\zeta)}\lfloor(2a)^{-1}T^{\zeta}\rfloor} \omega_{\mathsf{f}}(x+y, 2aT^{-K_1})\Psi^+_{T,\mathsf{p},\mathcal{C}}(dy)\Big\|_1. \qquad (56)
\end{aligned}$$

Here $\mathcal{K}$ is a probability measure on $\mathbb{R}^{\mathsf{d}}$ having the Fourier transform with a compact support, and $\mathcal{K}_\epsilon$ is defined by $\mathcal{K}_\epsilon(A) = \mathcal{K}(\epsilon^{-1}A)$for $A \in \mathbb{B}_{\mathsf{d}}$ and $\epsilon > 0$. K_1 and K are arbitrary positive constants such that $K_1 > K$, and $\zeta = (K_1 - K)/2$. A positive constnat $\alpha > 1/2$ and $\delta(\alpha) = \alpha^{-1}(1-\alpha) \in (0,1)$. C is a constant depending on $\mathsf{p}, \mathsf{d}, \alpha$, and $C' = (2\alpha - 1)^{-1}$, and a is a positive number. The function $h(x) = 1 + x^{\mathsf{p}_0}$. For the three terms in $\{...\}$ on the right-hand side of (56), the same estimate is valid as in Yoshida [18], p.605. The last term is estimated with

$$\Big\|\psi_T\Phi_T\omega\big(\mathsf{f}; T^{-K}, \Psi^+_{T,\mathsf{p},\mathcal{C}}\big)\Big\|_1$$

for large T. Thus, we obtained the following lemma, that is a counterpart to Lemma 11 of Yoshida [18], p.601.

Lemma 3. *Assume Conditions (41) and (43) are fulfilled. Then, for given constants K_1 and K satisfying $K_1 > K > 0$, there exist positive constants C_{d} and δ'' such that*

$$\begin{aligned}
&\left\| \psi_T \Big(E_{\mathcal{C}}\big[\mathsf{f}(T^{-1/2}Z_T)\big] - \Psi_{T,p,\mathcal{C}}[\mathsf{f}] \Big) \right\|_1 \\
\leq\ & C_{\mathsf{d}} M \sum_{\alpha:|\alpha|\leq \mathsf{d}+1+\mathsf{p}_0} \int \left\| \psi_T \partial_u^\alpha \big[\big(H_T(u,\mathcal{C}) e^{\mathbf{i}u\cdot e_T} - \widehat{\Psi}_{T,\mathsf{p},\mathcal{C}}(u) \big) \widehat{\mathcal{K}}(T^{-K_1}u) \big] \right\|_1 du \\
& + C_{\mathsf{d}} \left\| \psi_T \Phi_T \omega\big(\mathsf{f}; T^{-K}, \Psi^+_{T,\mathsf{p},\mathcal{C}}\big) \right\|_1 + \overline{o}_{\check{\mathcal{E}}(M,\gamma)}\big(T^{-(\mathsf{p}-2+\delta'')/2}\big)
\end{aligned} \tag{57}$$

as $T \to \infty$, that is, this inequality is valid for $\mathsf{f} \in \check{\mathcal{E}}(M,\gamma)$, having the error term of order $o\big(T^{-(\mathsf{p}-2+\delta'')/2}\big)$ uniformly in $\mathsf{f} \in \check{\mathcal{E}}(M,\gamma)$.

Proof of Theorem 5. Following Yoshida [18], we estimate the integral on the right-hand side of (57) by dividing it into two integrals over $\{u;\ |u| \geq l(T)\}$ (Lemma 10, p.594) and $\{u;\ |u| \leq l(T)\}$ (Lemma 7, p.591), $l(T) = \lfloor T^{\epsilon_4} \rfloor$ defined with a sufficiently small positive number ϵ_4 (Lemma 2, p.587). In Yoshida [18], the resulting estimates were combined in Lemma 8 (p.592, p.596) to give Proposition 1 (p.593), and Theorem 1 (p.565), that is rephrased in the present situation as Theorem 5. For the proof of Theorem 5, what is different in our situation is estimation of the second term on the right-hand side of (57); it is originally on p.600 of Yoshida [18]. As mentioned on p.600 of Yoshida [18], the measure $\Psi^+_{T,\mathsf{p},\mathcal{C}}$ has an expression

$$\phi\big(x; 0, \mathrm{Var}_{\mathcal{C}}[T^{-1/2}Z_T]\big) \Bigg(\sum_{\mathbf{n}:|\mathbf{n}|\leq 3(\mathsf{p}-2)} d_{\mathbf{n}}(\mathcal{C}) x^{\mathbf{n}} \Bigg)^+ \tag{58}$$

and the coefficients $d_{\mathbf{n}}(\mathcal{C})$ may involve the conditional cumulants of $T^{-1/2}Z_T$ of degree at most p, and $\big(\det \mathrm{Var}_{\mathcal{C}}[T^{-1/2}Z_T]\big)^{-k}$ $(k = 0, 1, ..., 3(\mathsf{p}-2))$. Therefore

$$\big\| \psi_T |d_{\mathbf{n}}(\mathcal{C})| \big\|_2 \quad \lesssim \quad s_T^{-3(\mathsf{p}-2)\mathsf{d}} \tag{59}$$

since the factors other than the determinants are in $L^{\infty-}$. On the other hand,

$$\begin{aligned}
&\left\| \psi_T \Phi_T |d_{\mathbf{n}}(\mathcal{C})| \int_{\mathbb{R}^{\mathsf{d}}} \omega_{\mathsf{f}}(x; T^{-K}) |x|^{|\mathbf{n}|} \phi\big(x; 0, \mathrm{Var}_{\mathcal{C}}[T^{-1/2}Z_T]\big) dx \right\|_1 \\
\leq\ & \big\| \psi_T |d_{\mathbf{n}}(\mathcal{C})| \big\|_2 \left\| \psi_T \Phi_T \int_{\mathbb{R}^{\mathsf{d}}} \omega_{\mathsf{f}}(x; T^{-K}) |x|^{|\mathbf{n}|} \phi\big(x; 0, \mathrm{Var}_{\mathcal{C}}[T^{-1/2}Z_T]\big) dx \right\|_2 \\
\leq\ & u_T^{\mathsf{d}/2} s_T^{-\mathsf{d}/2} \big\| \psi_T |d_{\mathbf{n}}(\mathcal{C})| \big\|_2 \left\| \psi_T \Phi_T \int_{\mathbb{R}^{\mathsf{d}}} \omega_{\mathsf{f}}(x; T^{-K}) |x|^{|\mathbf{n}|} \phi\big(x; 0, u_T I_{\mathsf{d}}\big) dx \right\|_2
\end{aligned} \tag{60}$$

since

$$\phi\big(x; 0, \mathrm{Var}_{\mathcal{C}}[T^{-1/2}Z_T]\big) \quad \leq \quad (\det s_T I_{\mathsf{d}})^{-1/2} (\det u_T I_{\mathsf{d}})^{1/2} \phi\big(x; 0, u_T I_{\mathsf{d}}\big)$$

when $\psi_T = 1$. Moreover,

$$\begin{aligned}
&\left\| \psi_T \Phi_T \int_{\mathbb{R}^{\mathsf{d}}} \omega_{\mathsf{f}}(x; T^{-K}) |x|^{|\mathbf{n}|} \phi\big(x; 0, u_T I_{\mathsf{d}}\big) dx \right\|_2 \\
\leq\ & \left\| \psi_T \Phi_T \Big(\int_{\mathbb{R}^{\mathsf{d}}} \omega_{\mathsf{f}}(x; T^{-K})^2 \phi\big(x; 0, u_T I_{\mathsf{d}}\big) dx \Big)^{1/2} \Big(\psi_T \int_{\mathbb{R}^{\mathsf{d}}} |x|^{2|\mathbf{n}|} \phi\big(x; 0, u_T I_{\mathsf{d}}\big) dx \Big)^{1/2} \right\|_2 \\
\leq\ & \left\| \Phi_T \Big(\int_{\mathbb{R}^{\mathsf{d}}} \omega_{\mathsf{f}}(x; T^{-K})^2 \phi\big(x; 0, u_T I_{\mathsf{d}}\big) dx \Big)^{1/2} \right\|_2 \left\| \Big(\int_{\mathbb{R}^{\mathsf{d}}} |x|^{2|\mathbf{n}|} \phi\big(x; 0, u_T I_{\mathsf{d}}\big) dx \Big)^{1/2} \right\|_\infty \\
\lesssim\ & u_T^{|\mathbf{n}|/2} \omega_2^{\Phi}\big(\mathsf{f}; T^{-K}, N(0, u_T I_{\mathsf{d}})\big)
\end{aligned} \tag{61}$$

From (58), (60) and (61), we obtain

$$\begin{aligned}
&\left\| \psi_T \Phi_T \omega\big(\mathsf{f}; T^{-K}, \Psi^{+}_{T,\mathsf{p},\mathcal{C}}\big) \right\|_1 \\
\leq \quad & \sum_{\mathbf{n}\in\mathbb{Z}^{\mathsf{d}}_{+}:|\mathbf{h}|\leq 3(\mathsf{p}-2)} \left\| \psi_T \Phi_T |d_{\mathbf{n}}(\mathcal{C})| \int_{\mathbb{R}^{\mathsf{d}}} \omega_{\mathsf{f}}(x; T^{-K}) |x|^{|\mathbf{n}|} \phi\big(x; 0, \mathrm{Var}_{\mathcal{C}}[T^{-1/2} Z_T]\big) dx \right\|_1 \\
\lesssim \quad & u_T^{\gamma(1)} s_T^{-\gamma(2)} \omega_2^{\Phi}\big(\mathsf{f}; T^{-K}, N(0, u_T I_{\mathsf{d}})\big).
\end{aligned}$$

This completes the sketch of the proof of Theorem 5.

B.2 Asymptotic expansion applied to the backward approximation scheme for the bandit algorithm

B.2.1 Sequentially partial mixing process

We go back to the model for bandits algorithms in Section A.1.1. The error bound of the backward approximation formula will be provided by the asymptotic expansion scheme under a partially mixing property, generalized in Section B.1.

All random variables are assumed to be defined on some common probability space $(\Omega, \mathcal{F}, P)$. To correctly apply the theory prepared so far, we begin with rebuilding an environment as follows.

[(1)]$\mathbb{S} = \{1, ..., S\}$, $\mathbb{J}^{\infty} = \mathbb{S} \times \mathbb{N}$, $\mathbb{J}^{\infty}_s = \{j = (s,i) \in \mathbb{J}^{\infty};\ i \in \mathbb{N}\}$. $c = (c_s)_{s\in\mathbb{S}}$, $c_s \in \mathfrak{C}_s$. $\mathfrak{C} = \prod_{s\in\mathbb{S}} \mathfrak{C}_s$. $(A_j)_{j=(s,i)\in\mathbb{J}^{\infty}}$ is a collection of doubly indexed random vectors. A_j is $\overline{k}_{s(j)}$-dimensional, where $s(j) = s$ for $j = (s,i)$. $\mathbf{A}^{\infty}_s = (A_j)_{j\in\mathbb{J}^{\infty},\ s(j)=s}$. $(\epsilon_j)_{(j=(s,i)\in\mathbb{J}^{\infty}}$ is a collection of doubly indexed r_s-dimensional random vectors. Given a measurable space $\mathcal{L}_s$ for $s \in \mathbb{S}$, L^n_s is an $\mathcal{L}_s$-valued random map measurable with respect to $\sigma[\mathbf{A}^{\infty}_s]$. For $s \in \mathbb{S}$, G^s_T is a $\mathsf{d}_s \times \mathsf{d}_s$ random matrix, and H^s_j is a $\mathsf{d}_s \times \mathsf{r}_s$ random matrix for $j \in \mathbb{J}^{\infty}_s$. Both are measurable with respect to $\sigma[\mathbf{A}^{\infty}_s]$. $\mathcal{G}^{\infty}_s = \sigma\big[(L^n_{s'})_{n\in\mathbb{N}}, (G^{s'}_T)_{T\in\mathbb{N}}, (H^{s'}_t)_{t\in\mathbb{N}}, (\epsilon_j)_{j\in\mathbb{J}^{\infty}_s};\ s' \leq s\big]$ for $s \in \mathbb{S}$, and $\mathcal{G}^{\infty}_0$ is the trivial σ-field. $\mathcal{C}^s = \mathcal{G}^{\infty}_{s-1} \vee \sigma[\mathbf{A}^{\infty}_s]\ \ (\epsilon_j)_{j\in\mathbb{J}^{\infty}_s}\mathcal{C}^s$ for every $s \in \mathbb{S}$. $W^n_j = n_s^{-1/2} G^s_{n_s} H^s_j$ for $n \in \mathbb{N}$ and $j \in \mathbb{J}^{\infty}_s$. $Z^s_T = \sum_{i=1}^{T} H^s_{(s,i)} \epsilon_{(s,i)}$ for $T \in \mathbb{N}$. In particular, $Z^s_0 = 0$. n_s $(s \in \mathbb{S})$ are positive integers depending on $n \in \mathbb{N}$ such that $\inf_{s\in\mathbb{S}} \liminf (n_s/n) > 0$. $\mathbb{Z}^n_s = \sum_{j\in\mathbb{J}^n_s} W^n_j \epsilon_j = G^s_{n_s} n_s^{-1/2} Z^s_{n_s}$, $\mathbb{J}^n_s = \{j = (s,i);\ i \leq n_s\}$. $\{\mathcal{B}^s_I\}_{I\subset\mathbb{R}_+}$ is a family of sub σ-fields of $\mathcal{F}$ such that $\mathcal{B}^s_{I_1} \subset \mathcal{B}^s_{I_2}$ for $I_1 \subset I_2 \subset \mathbb{R}_+$. $\epsilon_{s,i}$ is $\mathcal{B}^s_{[i]} \vee \mathcal{C}^s$-measurable for $s \in \mathbb{S}$ and $i \in \mathbb{N}$. $\alpha^s(h|\mathcal{C}^s) : \Omega \to [0,1]$ is a $\mathcal{C}^s$-measurable function for $h \in \mathbb{R}_+$ such that

$$\alpha^s(h|\mathcal{C}^s) \quad \geq \quad \sup_{\substack{h'\geq h \\ t\in\overline{\mathbb{R}}_+}} \ \sup_{\substack{B_1\in\mathcal{B}^s_{[0,t]}\vee\mathcal{C}^s \\ B_2\in\mathcal{B}^s_{[t,\infty)}\vee\mathcal{C}^s}} \big| P_{\mathcal{C}^s}[B_1 \cap B_2] - P_{\mathcal{C}^s}[B_1] P_{\mathcal{C}^s}[B_2] \big| \quad a.s.$$

Φ^T_s is a $\{0,1\}$-valued $\mathcal{C}^s$-measurable function for $(T,s) \in \mathbb{N} \times \mathbb{S} \times \mathfrak{C}$, and $\Phi^{n_s}_s \leq 1_{\{L^n_s \in \mathcal{A}^n_s\}}$ for a measurable set $\mathcal{A}^n_s \subset \mathcal{L}_s$ for $(n,s) \in \mathbb{N} \times \mathbb{S}$

$$\sup_{n\in\mathbb{N}, s\in\mathbb{S}, \omega\in\Omega} \big(\Phi^n_{s,c} |G^s_{n_s}| + \Phi^n_{s,c} |(G^s_{n_s})^{-1}| \big) \quad < \quad \infty.$$

Condition (9) is Condition (16). Under this condition, each $\mathcal{C}^s$-conditional cumulant is a constant not depending on $\mathcal{C}^s$. We assume the following conditions.

[B1]1 There exists a positive constant a such that

$$\big\| \alpha^s(h|\mathcal{C}^s) \big\|_1 \quad \leq \quad a^{-1} e^{-ah} \qquad (h > 0,\ s \in \mathbb{S}).$$

[B2] For every $L > 0$,

$$\max_{s\in\mathbb{S}} \sup_{j\in\mathbb{J}^{\infty}_s} E\Big[\big| E_{\mathcal{C}^s}[|H^s_j \epsilon_j|^{\mathsf{p}+1}] \big|^L \Big] \quad < \quad \infty$$

and $E_{\mathcal{C}^s}[H^s_j \epsilon_j] = 0$ for $j \in \mathbb{J}^{\infty}_s$, $s \in \mathbb{S}$.

For $s \in \mathbb{S}$, a collection of intervals $(I^s_{T,\ell})_{\ell\in\{1,...,n'(T)\},\, T\in\mathbb{N}}$ is called (dense) reduction intervals associated with $\mathcal{C}^s$ and $\left(\widehat{\mathcal{C}}^s_{T,\ell}\right)_{\ell\in\{1,...,n'(T)\},\, T\in\mathbb{N}}$ if the following conditions are satisfied.

[(i)]$\widehat{\mathcal{C}}^s_{T,\ell}$ ($\ell \in \{1,...,n'(T)\}$, $T \in \mathbb{N}$) are sub σ-fields of $\mathcal{F}$ with $\widehat{\mathcal{C}}^s_{T,\ell} \supset \mathcal{C}^s$. $I^s_{T,\ell} = [u^s_{T,\ell}, v^s_{T,\ell}]$ satisfies $v^s_{T,\ell} \le u^s_{T,\ell+1}$ for $\ell \in \{1,...,n'(T)-1$, $0 < \inf_{T,\ell}\left(v^s_{T,\ell} - u^s_{T,\ell}\right) \le \sup_{T,\ell}\left(v^s_{T,\ell} - u^s_{T,\ell}\right) < \infty$, and $\liminf_{T\to\infty}\left(n'(T)/T\right) > 0$. For $s \in \mathbb{S}$ and any subsequence $\ell_1,...,\ell_{n''(T)}$ of $1,...,n'(T)$, any bounded $\mathcal{B}_{[0,T]\setminus\cup_{\ell\in\{\ell_1,...,\ell_{n''(T)}\}} I^s_{T,\ell}} \vee \mathcal{C}^s$-measurable random variable A and any bounded $\mathcal{B}^s_{I^s_{T,\ell''}} \vee \mathcal{C}^s$-measurable random variable $\mathsf{A}_{\ell''}$, $\ell'' \in \{\ell_1,...,\ell_{n''(T)}\}$, it holds that

$$E_{\mathcal{C}^s}\left[\mathsf{A}\mathsf{A}_{\ell_1}\cdots\mathsf{A}_{\ell_{n''(T)}}\right] \quad = \quad E_{\mathcal{C}^s}\left[\mathsf{A}E_{\widehat{\mathcal{C}}^s_{T,\ell_1}}[\mathsf{A}_{\ell_1}]\cdots E_{\widehat{\mathcal{C}}^s_{T,\ell_{n''(T)}}}[\mathsf{A}_{\ell_{n''(T)}}]\right]. \tag{62}$$

The condition (62) is a conditional type of Markovian property. For example, if the underlying process of the system is Markovian, then we can take the σ-field $\widehat{\mathcal{C}}^s_{T,\ell}$ consisting of the information of the process at the boundary of $I^s_{T,\ell}$.

We consider the following condition for the classical Cramér condition.

[B3] For every $L > 0$, there exist measurable maps $\psi^s_{T,\ell} : \Omega \to [0,1]$ ($s \in \mathbb{S}$, $T \in \mathbb{N}$ and every $\ell \in \{1,....,n'(T)\}$), and there exist positive constants η_1, η_2, η_3 and B such that $\eta_1+\eta_2 < 1$, $\eta_3 < 1$ and that

$$P\left[\sum_{\ell=1}^{n'(T)} E_{\mathcal{C}^s}\left[\widetilde{\Phi}^s_{T,\ell} \le 1-\eta_2\right] > \eta_3 n'(T)\right] \quad = \quad O(T^{-L})$$

as $T \to \infty$, where

$$\widetilde{\Phi}^s_{T,\ell} \quad = \quad E_{\widehat{\mathcal{C}}^s_{T,\ell}}[\psi^s_{T,\ell}]\cdot 1_{\left\{\sup_{\substack{u\in\mathbb{R}^{\mathtt{d}_s}\\ |u|\ge B}}\left|E_{\widehat{\mathcal{C}}^s_{T,\ell}}\left[\psi^s_{T,\ell}\exp\left\{\mathtt{i}u\cdot\left(Z^s_{v^s_{T,\ell}}-Z^s_{u^s_{T,\ell}}\right)\right\}\right]\right|<\eta_1\right\}}.$$

Asymptotic expansion for a stochastic differential equation under a random environment is discussed by Yoshida [18]. Such condition is used to ensure regularity of the distribution of the targeted variable when a collective, conditional type Cramér condition $[B3]$ is satisfied. The Malliavin calculus worked effectively to verify it for stochastic processes. On the other hand, as we will see later, this condition becomes quite simple for a conditionally independent experiments.

class $\mathcal{D}^s$

$$r_{s-1}(\overline{z}_{s-1})f^n_{s-1}$$

[B4] **(i)** There exists a positive constant δ such that

$$\mathrm{Var}_{\mathcal{C}^s}[n_s^{-1/2}Z^s_{n_s}] > \delta I_{\mathtt{d}_s} \qquad a.s. \text{ on } \{L^n_s \in \mathcal{A}^n_{s,c}\}$$

for $(s,c,n) \in \mathbb{S}\times\mathfrak{C}\times\mathbb{N}$.

(ii)

$$\sup_{T\in\mathbb{N},s\in\mathbb{S},\omega\in\Omega}\left(\Phi^T_s|G^s_T| + \Phi^T_s|(G^s_T)^{-1}|\right) \quad < \quad \infty.$$

(iii) $P[\Phi^T_s = 0] = O(T^{-L})$ as $n \to \infty$ for any $L > 0$.

At batch s, we use the $\mathcal{C}^s$-measurable random signed measure $\Psi^n_{s,\mathsf{p},\mathbf{w}^n_s}$ in Section B. In the environment of this section, the asymptotic expansion $\Psi^*_{n_s,\mathsf{p},\mathcal{C}^s}$ for $\mathbb{Z}^n_s$ serves as $\Psi^n_{s,\mathsf{p},\mathbf{w}^n_s}$. We remark that $\Psi^*_{T,\mathsf{p},\mathcal{C}}$ was generically defined by (50) for Z^*_T of (46). More precisely, for

$$\chi_{n_s,r,\mathcal{C}^s}(u) \quad = \quad \left(\frac{\partial}{\partial\epsilon}\right)^r \log E_{\mathcal{C}^s}\left[\exp\left(\mathtt{i}\epsilon u\cdot n_s^{-1/2}\mathbb{Z}^n_s\right)\right]\Big|_{\epsilon=0} \tag{63}$$

for $u \in \mathbb{R}^{\mathsf{d}_s}$, the functions $P_{n_s,r,\mathcal{C}^s}\big(u;\chi_{n_s,\cdot,\mathcal{C}^s}\big)$ are determined by the formal expansion

$$\exp\left(\sum_{r=2}^{\infty}\epsilon^{r-2}(r!)^{-1}\chi_{n_s,r,\mathcal{C}^s}(u)\right) \quad = \quad \exp\big(2^{-1}\chi_{n_s,2,\mathcal{C}^s}(u)\big)\sum_{r=0}^{\infty}\epsilon^r n_s^{-r/2}P_{n_s,r,\mathcal{C}^s}\big(u;\chi_{n_s,\cdot,\mathcal{C}^s}\big) \tag{64}$$

We remark that $\chi_{n_s,r,\mathcal{C}^s}(u)$ is a function of $\mathbf{W}_s^n$ under Condition (9). Therefore, for each (n_s,r), there exists a measurable function $\overline{P}_{n_s,r,\mathbf{w}_s^n}(u)$ of $\mathbf{w}_s^n$ such that

$$P_{n_s,r,\mathcal{C}^s}\big(u;\chi_{n_s,\cdot,\mathcal{C}^s}\big) \quad = \quad \overline{P}_{n_s,r,\mathbf{w}_s^n}(u)\big|_{\mathbf{w}_s^n=\mathbf{W}_s^n}.$$

Define $\widehat{\Psi}^n_{s,\mathsf{p},\mathbf{w}_s^n}$ by

$$\widehat{\Psi}^n_{s,\mathsf{p},\mathbf{w}_s^n}(u) \quad = \quad \exp\big(2^{-1}\chi_{n_s,2,\mathcal{C}^s}(u)\big)\sum_{r=0}^{\mathsf{p}-2}n_s^{-r/2}\overline{P}_{n_s,r,\mathbf{w}_s^n}(u) \tag{65}$$

and the random signed measure $\Psi^n_{s,\mathsf{p},\mathbf{w}_s^n}$ on $\mathbb{R}^{\mathsf{d}_s}$ is defined by the Fourier inversion of $\widehat{\Psi}^n_{s,\mathsf{p},\mathbf{w}_s^n}$:

$$\Psi^n_{s,\mathsf{p},\mathbf{w}_s^n} \quad = \quad \mathcal{F}^{-1}\big[\widehat{\Psi}^n_{s,\mathsf{p},\mathbf{w}_s^n}\big]. \tag{66}$$

In other words, $\Psi^n_{s,\mathsf{p},\mathbf{w}_s^n}$ is absolutely continuous on the event $\mathsf{E}_{n_s,s}=\{\chi_{n_s,2,\mathcal{C}^s}$ is non-degenerate$\}$, and the local density on it is

$$\frac{d\Psi^n_{s,\mathsf{p},\mathbf{w}_s^n}}{dz}(z) \quad = \quad \sum_{r=0}^{\mathsf{p}-2}n_s^{-r/2}\overline{P}_{n_s,r,\mathbf{w}_s^n}(-\mathtt{i}\partial_z)^{\phi}\big(z;0,\mathrm{Var}_{\mathcal{C}^s}[\mathbb{Z}_s^n]\big) \tag{67}$$

for $z\in\mathbb{R}^{\mathsf{d}_s}$, where makes the adjoint operator of $\overline{P}_{n_s,r,\mathbf{w}_s^n}(-\mathtt{i}\partial_z)$. Specifying $\Psi^n_{s,\mathsf{p},\mathbf{w}_s^n}$ outside of $\mathsf{E}_{n_s,s}$ is not important but we may set $\Psi^n_{s,\mathsf{p},\mathbf{w}_s^n}=\delta_0$, the delta measure.

A representation of the local density of $\Psi^n_{s,\mathsf{p},\mathbf{w}_s^n}$ is presented by (67). We write

$$\lambda_{n_s,r,\mathcal{C}^s}[(\mathtt{i}u)^{\otimes r}] \quad = \quad n_s^{(r-2)/2}\chi_{n_s,r,\mathcal{C}^s}(u) \qquad (u\in\mathbb{R}^{\mathsf{d}_s}),$$

For notational simplicity, hereafter we abuse the symbol $\chi_{n_s,r,\mathcal{C}^s}(u)$ for its representation as a function of $\mathbf{w}_s^n$, like $\overline{P}_{n_s,r,\mathbf{w}_s^n}(u)$ for $P_{n_s,r,\mathcal{C}^s}\big(u;\chi_{n_s,\cdot,\mathcal{C}^s}\big)$. This rule also applies to $\lambda_{n_s,r,\mathcal{C}^s}$. Then the formula (67) gives a local density representation of $\Psi^n_{s,\mathsf{p},\mathbf{w}_s^n}$ as

$$\begin{aligned}\frac{d\Psi^n_{s,\mathsf{p},\mathbf{w}_s^n}}{dz}(z) \quad = \quad & \Big\{1+\frac{1}{6}n_s^{-1/2}\lambda_{n_s,3,\mathcal{C}^s}[(-\partial_z)^{\otimes 3}] \\ & +n_s^{-1}\left(\frac{1}{24}\lambda_{n_s,4,\mathcal{C}^s}[(-\partial_z)^{\otimes 4}]+\frac{1}{72}(\lambda_{n_s,3,\mathcal{C}^s}[(-\partial_z)^{\otimes 3}])^2\right)+\cdots\Big\} \\ & \times\phi\big(z;0,\mathrm{Var}_{\mathcal{C}^s}[\mathbb{Z}_s^n]\big).\end{aligned} \tag{68}$$

With the tensors consisting of the ${\mathsf{d}_s}^r$ Hermite polynomials

$$h_r(z;\Sigma) \quad = \quad (-1)^r\phi(z;0,\Sigma)^{-1}\partial_z^r\phi(z;0,\Sigma),$$

we can rewrite (68) as

$$\begin{aligned}\frac{d\Psi^n_{s,\mathsf{p},\mathbf{w}_s^n}}{dz}(z) \quad = \quad & \Big\{1+\frac{1}{6}n_s^{-1/2}\lambda_{n_s,3,\mathcal{C}^s}\big[h_3(z;\mathrm{Var}_{\mathcal{C}^s}[\mathbb{Z}_s^n])\big] \\ & +n_s^{-1}\left(\frac{1}{24}\lambda_{n_s,4,\mathcal{C}^s}\big[h_4(z;\mathrm{Var}_{\mathcal{C}^s}[\mathbb{Z}_s^n])\big]+\frac{1}{72}(\lambda_{n_s,3,\mathcal{C}^s})^{\otimes 2}\big[h_6(z;\mathrm{Var}_{\mathcal{C}^s}[\mathbb{Z}_s^n])\big]\right) \\ & +\cdots\cdots\Big\}\phi\big(z;0,\mathrm{Var}_{\mathcal{C}^s}[\mathbb{Z}_s^n]\big).\end{aligned} \tag{69}$$

We once again remark that the $\mathcal{C}^s$-conditional cumulants and hence the right-hand side is a measurable function of $\mathbf{w}_s^n$.

Under $[B2]$ and $[B4]$ (i) (ii), the condition (33) is verified. Therefore the functions $f^n_{s-1}(\underline{l}_{s-1}, \underline{z}_{s-1})$ are inductively well defined. We consider the class $\mathcal{D}(\mathbf{M}, \gamma)$ of random functions. Recall this class satisfies the condition (32). Thus, at each batch of the backward scheme, we obtain an error bound taking the form of (45) with Z^s_T in place of Z_T, and hence of (52), thanks to Theorem 6. In each step $s \in \mathbb{S}$, let $\theta = 1/2$ e.g., and fix sequences $\mathsf{U}^s_{n_s}$ and $\mathsf{S}^s_{n_s}$ for "u_T" and "s_T", respectively. Then the component $(P[....])^\theta$ on the right-hand side of (52) is negligible since it is $O(n_s^{-(\mathsf{p}-1)/2})$, which follows from $[B4]$ (i) and (iii).

Assume that

$$\limsup_{T\to\infty} \Big\| \tag{70}$$

We restrict $\mathcal{D}(\mathbf{M}, \gamma)$ to $\widehat{\mathcal{D}}(M, \gamma)$ so that for some positive constant ϵ_0,

$$\sup_{\mathsf{f}\in\widehat{\mathcal{D}}(M,\gamma)} \sup_{n\in\mathbb{N}} \sup_{\underline{l}_s, \underline{z}_{s-1}} \omega_2^{\Phi_s^{'}}\big(f^n_s(\underline{l}_s, \underline{z}_{s-1}, \cdot); r, N(0, \mathsf{U}^s_{n_s} I_{\mathsf{d}_s})\big) \quad = \quad O(r^{\epsilon_0}) \qquad (r \downarrow 0) \tag{71}$$

for every $s \in \mathbb{S}$ and $c \in \mathfrak{C}$. We remember, under (71), it is possible to choose K, $\mathsf{U}^s_{n_s}$ and $\mathsf{S}^s_{n_s}$ to make

$$\big(\mathsf{U}^s_{n_s}\big)^{\gamma(1)}\big(\mathsf{S}^s_{n_s}\big)^{-\gamma(2)} \sup_{f\in\widehat{\mathcal{D}}(\mathbf{M},\gamma)} \omega_2^{\Phi_s^{'}}\big(f^n_s(\underline{L}^n_s, \underline{\mathbb{Z}}^n_{s-1}, \cdot); n^{-K/2}, N(0, \mathsf{U}^s_{n_s} I_{\mathsf{d}_s})\big) \quad = \quad o(n^{-(\mathsf{p}-2+\delta_1)/2})$$

as $n \to \infty$. See Remark 4. Then, applying Theorem 6 repeatedly, we obtain

$$\sup_{\mathsf{f}\in\widehat{\mathcal{D}}(M,\gamma)} \varrho^n_s(\mathsf{f}) \quad = \quad o(n^{-(\mathsf{p}-2+\delta_1)/2}) \tag{72}$$

as $n \to \infty$, where δ_1 is a positive constant. We should remark that here we repeatedly apply Theorem 6 at each batch s to the sequence $\big(\mathbb{Z}^n_s\big)_{n\in\mathbb{N}}$ of (13) made from the complete sequence $\big(Z^s_T\big)_{T\in\mathbb{N}}$ (or the sequence embedded into $T \in \mathbb{R}_+$) of (11). In order to satisfy the condition (41), we are assuming, for every $s \in \mathbb{S}$,

$$\limsup_{T\in\mathbb{N}, T\to\infty} \big\| |\Psi_{T,\mathsf{p},\mathcal{C}^s}|[1 + |\cdot|^{\mathsf{p}_0}]\big\|_q \quad < \quad \infty \tag{73}$$

for some $q > 1$, where $\Psi_{T,\mathsf{p},\mathcal{C}^s}$ is the signed measure of asymptotic expansion constructed for Z^s_T given in (11). By the above consideration, under (73), we obtain the following result, since the elements appearing in the error bound are well controlled along the subsequence (n_s) in $\mathbb{N}$, thanks to the assumptions.

Theorem 7. *Suppose that Conditions $[B1]$-$[B4]$ are fulfilled. Let $M_s > 0$ for $s \in \mathbb{S}$, and $\gamma > 0$. Then there exists a constant M^* such that $\mathbf{E}_{n,c} = f^n_{0,c}$ admits the following estimates:*

$$|f^n_{0,c} - \widehat{f}^n_{0,c}| \quad \le \quad M^*\Big(n^{-(\mathsf{p}-2+\delta_1)/2} \sum_{s\in\mathbb{S}} \prod_{i=2}^{s} U^n_i\Big) \tag{74}$$

for all $f \in \widehat{\mathcal{D}}(\mathbf{M}, \gamma)$, $\mathbf{M} = (M_s)_{s\in\mathbb{S}}$.

Theorem 7 suggests use of $\widehat{f}^n_{0,c}$ to approximate $\mathbf{E}_{n,c}$. The U^n_{s+1} is given in (34). Usually U^n_{s+1} is uniformly bounded in n or grows so slowly as to be controlled by the factor $n^{-\delta_1/2}$ in the error bound of (74). It depends on how to cut off a bad event by the truncation function $1_{\{L^n_s \in \mathcal{A}^n_{s,c}\}}$.

Integrating what we obtained so far, for the functionals $\mathbb{Y}^n_s = Y^n_s(L^n_s, \mathbb{Z}^n_s)$ of (20), we can approximate the expectation $\overline{\mathbf{E}}_n = E\big[f\big(\underline{L}^n_S, \underline{\mathbb{Y}}^n_S\big)\big]$ at (22) by using the expectations $\mathbf{E}_n$ at (25), based on Theorem 3. As shown in Theorem 7, the expectation $\mathbf{E}_n = f^n_0$ is approximated by $\widehat{f}^n_0$ defined by the backward scheme in (**??**):

$$\widehat{f}^n_{s-1}(\underline{l}_{s-1}, \underline{z}_{s-1}) \quad = \quad \int \widehat{f}^n_s\big(\underline{l}_s, \underline{z}_s\big)\Psi^n_{s,\mathsf{p},\mathbf{w}^n_s}(dz_s) 1_{\{l_s\in\mathcal{A}^n_{s,c}\}} \nu^n_{c_s}(dl_s, d\mathbf{w}^n_s) q\big((\underline{l}_{s-1}, \underline{z}_{s-1}), dc_s\big) \tag{75}$$

for $s \in \mathbb{S}$, where $\Psi_{n_s,p,\mathcal{C}^s}$ is the random signed measure constructed along (63)-(66) for $\mathbb{Z}^n_s$, having the representation (69).

Remark 6. Condition (33) is satisfied by $[B2]$ and $[B4]$.

B.2.2 Sequentially conditionally independent experiments

We consider the environment specified by (1)-(13) and (17) in Section B.2.1. Additionally, suppose that the sequence $(\epsilon_j)_{j\in\mathbb{J}_s^\infty}$ is a $\mathcal{C}^s$-conditionally independent and satisfies

$$\sup_{j\in\mathbb{J}_s^\infty} E_{\mathcal{C}^s}[|\epsilon_j|^{\mathsf{p}+1}] < \infty \text{ and } E_{\mathcal{C}^s}[\epsilon_j] \;=\; 0 \qquad (j\in\mathbb{J}_s^\infty) \tag{76}$$

for every $s\in\mathbb{S}$.

The intervals $I_{T,\ell}^s = [\ell-1,\ell]$ form dense reduction intervals. Condition $[B2]$ is rephrased as

[B2′] For every $L>0$,

$$\max_{s\in\mathbb{S}}\sup_{j\in\mathbb{J}_s^\infty} E\big[|H_j^s|^L\big] \quad < \quad \infty.$$

We replace $[B3]$ by a Cramér condition under the conditional probability:

[B3′] There exists a constant B_0 such that

$$\inf_{u\in\mathbb{R}^{\mathsf{d}_s}:|u|\geq B_0} \big|E_{\mathcal{C}^s}[\exp(\mathtt{i}u\cdot\epsilon_{(s,1)})]\big| \quad < \quad 1$$

for every $s\in\mathbb{S}$.

What is equivalent to Condition $[B4]$ (i) is the following: there exists a positive constant δ such that

$$n_s^{-1}\sum_{j\in\mathbb{J}_s^n} H_j^s(H_j^s)^\star \geq \delta I_{\mathsf{d}_s} \qquad a.s. \text{ on } \{L_s^n\in\mathcal{A}_{s,c}^n\}$$

for $(s,c,n)\in\mathbb{S}\times\mathfrak{C}\times\mathbb{N}$.

Take $\mathcal{B}_I^s$ so that $\mathcal{B}_{[i]}^s = \sigma[\epsilon_{(s,i)}]$. A choice of $\psi_{T,\ell}^s$ is

$$\psi_{T,\ell}^s \quad = \quad 1_{\left\{\lambda_{min}\left((G_T^sH_j^s)^\star G_T^sH_j^s\right)\geq\varepsilon_1\right\}}, \qquad j=(s,\ell),$$

for some positive constant ε_1, where λ_{min} denotes the minimum eigenvalue of a symmetric matrix. Now we apply Theorem 7 to obtain

Theorem 8. *Suppose that $(\epsilon_j)_{j\in\mathbb{J}_s^\infty}$ is a $\mathcal{C}^s$-conditionally independent sequence satisfying (76) and that Conditions $[B2']$, $[B3']$ and $[B4]$ are fulfilled. Then the estimate (74) is valid for all $f\in\widehat{\mathcal{D}}(\mathbf{M},\gamma)$, for some constant M^* for given $\mathbf{M}=(M_s)_{s\in\mathbb{S}}\in(0,\infty)^S$ and $\gamma>0$.*

C Linear and non-linear statistics of hierarchically conditionally i.i.d. sequences

We consider the setting in Section B.2.2, that is, additionally to Condition (9) of Section B.2.1 (originally presented as (16)), we assume that given $\mathcal{G}_{s-1}^\infty$, $(\epsilon_j)_{j\in\mathbb{J}_s^\infty}$ is a $\mathcal{C}^s$-conditionally independent and identically distributed sequence. An example of $(A_j)_{j\in\mathbb{J}_s^\infty}$ is a $\mathcal{G}_{s-1}^\infty$-conditionally independent and identically distributed sequence with $\mathcal{L}\{A_j|\mathcal{G}_{s-1}^\infty\} \;=\; \text{Multinomial}(1,\pi_s)$ for $j\in\mathbb{J}_s^n$, where $\pi_s\in\{p=(p_k)\in(0,1)^{\overline{k}_s};\sum_{k\in\mathcal{K}_s}p_k=1\}$. However, any distribution of $(A_j)_{j\in\mathbb{J}_s^\infty}$ is applicable in general. Even $\mathcal{G}_{s-1}^\infty$-conditional independency between $\{A_j\}_{j\in\mathbb{J}_s^\infty}$ is not necessary. So it is possible to treat an assignment mechanism that keeps at least a fixed number of observations for each arm. Suppose that

$$E_{\mathcal{C}^s}[|\epsilon_j|^{\mathsf{p}+1}] = \beta_{s,\mathsf{p}+1},\; E_{\mathcal{C}^s}[\epsilon_j] \;=\; 0 \text{ and } \mathrm{Var}_{\mathcal{C}^s}[\epsilon_j] \;=\; \sigma_s^2 \qquad (j\in\mathbb{J}_s^\infty), \tag{77}$$

where $\beta_{s,\mathsf{p}+1}$ is a finite constant and σ_s is a positive constant, for each $s\in\mathbb{S}$. Constancy of these conditional moments takes place since $(\epsilon_j)_{j\in\mathbb{J}_s^\infty}$ is independent of $\mathcal{C}^s$ (Condition (9) of Section B.2.1).

In particular, the $\mathcal{C}^s$-conditional cumulants of ϵ_j up to the order $\mathsf{p}+1$ are constants independent of $\mathcal{C}^s$. though we admit more complicated cases, e.g. the higher-order of moments depend on $\mathcal{C}^s$. It is also possible to relax the condition (77) as previously treated in the general setting.

Let $\mathcal{W}_s^n = \mathcal{G}_{t-1}^n \vee \sigma[\mathbf{A}_t^n]$ and $\mathcal{C}^s = \vee_{n\in\mathbb{N},t\in\mathbb{T}}\mathcal{W}_s^n$. Recall that we are assuming that the σ-fields $\mathcal{W}_s^n$ ($n\in\mathbb{N}$) have been constructed so that they are independent. The expansion of the $\mathcal{G}_t$-conditional distribution of

$$\mathbb{Z}_s^n \quad = \quad \sum_{j\in\mathbb{J}_s^n} W_j^n \epsilon_j$$

of (11) is given as follows. The $\mathcal{C}^s$-conditional cumulant functions $\chi_{n_s,r,\mathcal{C}^s}$ of $\mathbb{Z}_s^n$ are defined by

$$\chi_{n_s,r,\mathcal{C}^s}(u) \quad = \quad (\partial_\eta)_0^r \log E\big[\exp(\mathtt{i}\eta u_t\cdot\mathbb{Z}_T^n)\big]$$

for $u_t\in\mathbb{R}^{\mathsf{d}_t}$. The function $\widetilde{P}_{n_s,r,\mathcal{C}^s}$ is defined by the formal expansion

$$\exp\left(\sum_{r=2}^{\infty}\frac{1}{r!}\eta^{r-2}\chi_{n_s,r,\mathcal{C}^s}(u)\right) \quad = \quad \exp\left(\frac{1}{2}\chi_{n_s,2,\mathcal{C}^s}(u)\right)+\sum_{r=1}^{\infty}\eta^r n_s^{-r/2}\widetilde{P}_{n_s,r,\mathcal{C}^s}(u). \quad (78)$$

Then we make the expansion formula $\Psi_{n_s,rp\mathcal{C}^s}$ by

$$\Psi_{n_s,r,\mathcal{C}^s} \quad = \quad \mathcal{F}^{-1}\big[\widehat{\Psi}_{n_s,rp\mathcal{C}^s}\big],$$

the Fourier inversion of $\widehat{\Psi}_{n_s,p,\mathcal{C}^s}$ given by

$$\widehat{\Psi}_{n_s,p,\mathcal{C}^s}(u) \quad = \quad \exp\left(\frac{1}{2}\chi_{n_s,2,\mathcal{C}^s}(u)\right)+\sum_{r=1}^{p-2}n_s^{-r/2}\widetilde{P}_{n_s,r,\mathcal{C}^s}(u).$$

C.0.1 Asymptotic expansion for a linear functional

[0.3] We consider the setting in Section **??**, that is, additionally to Condition (16), we assume that given $\mathcal{G}_{t-1}^n$, each of $(\epsilon_j)_{j\in\mathbb{J}_s^n}$ and $(A_j)_{j\in\mathbb{J}_s^n}$ is an independent and identically distributed sequence, and $\mathcal{L}\{A_j|\mathcal{G}_{t-1}^n\} = \text{Multinomial}(1,\pi_t)$ for $j\in\mathbb{J}_s^n$, except for the conditional Gaussianity of $\mathcal{L}\{\epsilon_j|\mathcal{G}_{t-1}^n\}$ for $j\in\mathbb{J}_s^n$.

Let $\mathcal{W}_s^n = \mathcal{G}_{t-1}^n \vee \sigma[\mathbf{A}_t^n]$ and $\mathcal{C}^s = \vee_{n\in\mathbb{N},t\in\mathbb{T}}\mathcal{W}_s^n$. Recall that we are assuming that the σ-fields $\mathcal{W}_s^n$ ($n\in\mathbb{N}$) have been constructed so that they are independent. The expansion of the $\mathcal{G}_t$-conditional distribution of

$$\mathbb{Z}_s^n \quad = \quad \sum_{j\in\mathbb{J}_s^n} W_j^n \epsilon_j$$

of (11) is given as follows. The $\mathcal{C}^s$-conditional cumulant functions $\chi_{n_s,r,\mathcal{C}^s}$ of $\mathbb{Z}_s^n$ are defined by

$$\chi_{n_s,r,\mathcal{C}^s}(u) \quad = \quad (\partial_\eta)_0^r \log E\big[\exp(\mathtt{i}\eta u_t\cdot\mathbb{Z}_T)\big]$$

for $u_t\in\mathbb{R}^{\mathsf{d}_t}$. The function $\widetilde{P}_{n_s,r,\mathcal{C}^s}$ is defined by the formal expansion

$$\exp\left(\sum_{r=2}^{\infty}\frac{1}{r!}\eta^{r-2}\chi_{n_s,r,\mathcal{C}^s}(u)\right) \quad = \quad \exp\left(\frac{1}{2}\chi_{n_s,2,\mathcal{C}^s}(u)\right)+\sum_{r=1}^{\infty}\eta^r n_s^{-r/2}\widetilde{P}_{n_s,r,\mathcal{C}^s}(u). \quad (79)$$

Then we make the expansion formula $\Psi_{n_s,\mathsf{p},\mathcal{C}^s}$ by

$$\Psi_{n_s,\mathsf{p},\mathcal{C}^s} \quad = \quad \mathcal{F}^{-1}\big[\widehat{\Psi}_{n_s,\mathsf{p},\mathcal{C}^s}\big],$$

the Fourier inversion of $\widehat{\Psi}_{n_s,sfp,calc^s}$ given by

$$\widehat{\Psi}_{n_s,\mathsf{p},\mathcal{C}^s}(u) \quad = \quad \exp\left(\frac{1}{2}\chi_{n_s,2,\mathcal{C}^s}(u)\right)+\sum_{r=1}^{p-2}n_s^{-r/2}\widetilde{P}_{n_s,r,\mathcal{C}^s}(u).$$

Let $N_s^n = \big(N_{s,k_s}^n\big)_{k_s\in\mathcal{K}_s}$ with $N_{s,k_s}^n = \sum_{j\in\mathbb{J}_s^n}A_{j,k_s}$ for $k_s\in\mathcal{K}_s$, $s\in\mathbb{S}$. Suppose that

$$W_j^n = w_{n,s}(N_s^n/n_s)A_j, \quad (80)$$

where $w_{n,s} : [0,1]^{\overline{k}_s} \to \mathbb{R}^{\overline{k}_s} \otimes \mathbb{R}^{\overline{k}_s}$, the set of $\overline{k}_t \times \overline{k}_s$-matrices, is a measurable map, and the function $w_{n,t}$ is given by

$$w_{n,t}(x_1, ..., x_{\overline{k}_s}) \quad = \quad \mathrm{diag}\big(g_1(x_1)x_1^{-1/2}{n_s}^{-1/2}, ..., g_{\overline{k}_s}(x_{\overline{k}_s})x_{\overline{k}_s}^{-1/2}{n_s}^{-1/2}\big) \tag{81}$$

for some positive measurable functions $g_1, ..., g_{\overline{k}_s}$. Moreover, we assume that each $w_{n,s}(N_s^n/n_s)$ is invertible whenever $L_s^n \in \mathcal{A}_{s,c}^n$. In Section C.0.1, ϵ_j is one-dimensional.

We have

$$\begin{aligned}
\mathbb{Z}_s^n &= \sum_{j \in \mathbb{J}_s^n} W_j^n \epsilon_j \\
&= \mathrm{diag}\Bigg(g_1(N_{s,1}^n/n_s)(N_{s,1}^n)^{-1/2} \sum_{j \in \mathbb{J}_s^n} A_{j,1}\epsilon_j, ..., g_{\overline{k}_s}(N_{s,\overline{k}_s}^n/n_s)(N_{s,\overline{k}_s}^n)^{-1/2} \sum_{j \in \mathbb{J}_s^n} A_{j,\overline{k}_s}\epsilon_j \Bigg).
\end{aligned}$$

By the conditional independency of $\{\epsilon_j\}_{j \in \mathbb{J}_s^n}$,

$$\begin{aligned}
E\big[\exp(\mathtt{i}\eta u_s \cdot \mathbb{Z}_s^n)\big] &= E_{\mathcal{C}^s}\Bigg[\exp\Bigg(\sum_{k \in \mathcal{K}_s} \eta u_k g_k(N_{s,k}^n/n_s)(N_{s,k}^n)^{-1/2} \sum_{j \in \mathbb{J}_s^n} A_{j,k}\epsilon_j\Bigg)\Bigg] \\
&= \prod_{j \in \mathbb{J}_s^n} E_{\mathcal{C}^s}\Bigg[\exp\Bigg(\eta \sum_{k \in \mathcal{K}_s} u_k g_k(N_{s,k}^n/n_s)(N_{s,k}^n)^{-1/2} A_{j,k}\epsilon_j\Bigg)\Bigg],
\end{aligned}$$

and hence

$$\begin{aligned}
\chi_{n_s,r,\mathcal{C}^s}(u) &= (\partial_\eta)_0^r \log E\big[\exp(\mathtt{i}\eta u_s \cdot \mathbb{Z}_s^{n_s})\big] \\
&= \sum_{j \in \mathbb{J}_s^n} \mathtt{i}^r \kappa_{r,\mathcal{C}^s}\Bigg(\sum_{k \in \mathcal{K}_s} u_k g_k(N_{s,k}^n/n_s)(N_{s,k}^n)^{-1/2} A_{j,k}\epsilon_j\Bigg) \\
&\qquad\qquad (\kappa_{r,\mathcal{C}^s} : \text{ the } r\text{-th } \mathcal{C}^s\text{-conditional cumulant}) \\
&= \sum_{j \in \mathbb{J}_s^n} \mathtt{i}^r \Bigg(\sum_{k \in \mathcal{K}_s} u_k g_k(N_{s,k}^n/n_s)(N_{s,k}^n)^{-1/2} A_{j,k}\Bigg)^r \kappa_{r,\mathcal{C}^s}(\epsilon_{(s,1)}) \\
&= \sum_{j \in \mathbb{J}_s^n} \sum_{k \in \mathcal{K}_s} (\mathtt{i}u_k)^r \big\{g_k(N_{s,k}^n/n_s)\big\}^r (N_{s,k}^n)^{-r/2} A_{j,k}\, \kappa_{r,\mathcal{C}^s}(\epsilon_{(s,1)}) \\
&= \sum_{k \in \mathcal{K}_s} (\mathtt{i}u_k)^r \big\{g_k(N_{s,k}^n/n_s)\big\}^r (N_{s,k}^n)^{-(r-2)/2}\, \kappa_{r,\mathcal{C}^s}(\epsilon_{(s,1)}) \\
&= \sum_{k \in \mathcal{K}_s} n_s^{-(r-2)/2} \lambda_{n,k,r} (\mathtt{i}u_k)^r,
\end{aligned}$$

where the coefficients

$$\lambda_{n_s,r,\mathcal{C}^s,k} \quad = \quad n_s^{(r-2)/2} \cdot \big\{g_k(N_{s,k}^n/n_s)\big\}^r (N_{s,k}^n/n_s)^{-(r-2)/2}\, \kappa_{r,\mathcal{C}^s}(\epsilon_{(s,1)})$$

are random. Remark that the conditional variance of $\mathbb{Z}_s^{n_s}$ is a diagonal matrix due to the orthogonality between $(A_{j,k})_{k \in \mathcal{K}_s}$.

For example, if the functions g_{k_s} are identically equals to 1, as in the case of the batched OLS by Zhang et al. [20], we obtain

$$\begin{aligned}
\chi_{n_s,2,\mathcal{C}^s}(u) &= \sum_{k \in \mathcal{K}_s} (\mathtt{i}u_k)^2 \kappa_{2,\mathcal{C}^s}(\epsilon_{(s,1)}) \\
\chi_{n_s,r,\mathcal{C}^s}(u) &= \sum_{k \in \mathcal{K}_s} (\mathtt{i}u_k)^r (N_{s,k}^n)^{-(r-2)/2}\, \kappa_{r,\mathcal{C}^s}(\epsilon_{(s,1)}) \quad (r \geq 3),
\end{aligned}$$

and in particular,

$$\lambda_{n_s,2,\mathcal{C}^s,k} \quad = \quad \kappa_{2,\mathcal{C}^s}(\epsilon_{(s,1)}) \;=\; \sigma_s^2.$$

For the general functions g_k, the formula (69) shows the function $d\Psi_{n_s,p,\mathcal{C}^s}/dz(z)$ $(z = (z_1, ..., z_{\overline{k}_s}))$ is

$$\prod_{k\in\mathcal{K}_s} \phi(z_k;\lambda_{n_s,2,\mathcal{C}^s,k}) \times \prod_{k\in\mathcal{K}_s} \Big\{1 + \frac{1}{6} n_s^{-1/2} \lambda_{n_s,3,\mathcal{C}^s,k} h_3(z_k;\lambda_{n_s,2,\mathcal{C}^s,k}) + n_s^{-1}\Big(\frac{1}{24}\lambda_{n,k,4} h_4(z_k;\lambda_{n,k,2}) + \frac{1}{72}\lambda_{n,k,3}^2 h_6(z_k;\lambda_{n,k,2})\Big) + \cdots\Big\}, \tag{82}$$

the summation taken up to order $n_s^{-(p-2)/2}$.

C.0.2 Asymptotic expansion for a normalized estimator

In Section C.0.2, we assume that W_j^n takes the form

$$W_j^n = \begin{pmatrix} w_{n,s}(N_s^n/n_s) n_s^{-1/2} A_j & \mathbf{0} \\ 0 & n_s^{-1/2} \end{pmatrix} \qquad (j \in \mathbb{J}_s^n) \tag{83}$$

where $w_{n,s} : [0,1]^{\overline{k}_s} \to \mathbb{R}^{\overline{k}_s} \otimes \mathbb{R}^{\overline{k}_s}$, the set of $\overline{k}_s \times \overline{k}_s$-matrices, is a measurable map, and the function $w_{n,t}$ is given by (81). As before, $N_s^n = \big(N_{s,k_s}^n\big)_{k_s\in\mathcal{K}_s}$ with $N_{s,k_s}^n = \sum_{j\in\mathbb{J}_s^n} A_{j,k_s}$ for $k_s \in \mathcal{K}_s$, $s \in \mathbb{S}$. We assume that each $w_{n,s}(N_s^n/n_s)$ is invertible whenever $L_s^n \in \mathcal{A}_{s,c}^n$. We consider the error term ϵ_j set by

$$\epsilon_j = \begin{pmatrix} \dot{\epsilon}_j \\ \dot{\epsilon}_j^2 - \sigma_s^2 \end{pmatrix}$$

when $j \in \mathbb{J}_s^\infty$, where $\dot{\epsilon}_j$ is a one-dimensional random variable. Then

$$\mathbb{Z}_s^n = \sum_{j\in\mathbb{J}_s^n} W_j^n \epsilon_j = (\dot{\mathbb{Z}}_s^n, \ddot{\mathbb{Z}}_s^n) \tag{84}$$

for

$$\ddot{\mathbb{Z}}_s^n = n_s^{-1/2} \sum_{j\in\mathbb{J}_s^n} (\dot{\epsilon}_j^2 - \sigma_s^2). \tag{85}$$

By the conditional independency of $\{\epsilon_j\}_{j\in\mathbb{J}_s^n}$,

$$\begin{aligned} & E_{\mathcal{C}^s}\big[\exp(\mathtt{i}\eta u\cdot\mathbb{Z}_s^n)\big] \\ = \; & E_{\mathcal{C}^s}\Big[\exp\Big(\sum_{k\in\mathcal{K}_s} \eta u_k g_k(N_{s,k}^n/n_s)(N_{s,k}^n)^{-1/2} \sum_{j\in\mathbb{J}_s^n} A_{j,k}\dot{\epsilon}_j + \eta u_0\Big)\Big] \\ = \; & andhenceto \\ \chi_{n_s,r,\mathcal{C}^s}(u) = \; & (\partial_\eta)_0^r \log E_{\mathcal{C}^s}\big[\exp(\mathtt{i}\eta u\cdot\mathbb{Z}_s^n)\big] \\ = \; & \sum_{j\in\mathbb{J}_s^n} \mathtt{i}^r \kappa_{\mathcal{C}^s,r}\Big(\sum_{k\in\mathcal{K}_s} u_k g_k(N_{s,k}^n/n_s)(N_{s,k}^n)^{-1/2} A_{j,k}\dot{\epsilon}_j + u_0 n_s^{-1/2}(\dot{\epsilon}_j^2 - \sigma_s^2)\Big) \\ & \qquad (\kappa_{\mathcal{C}^s,r} : \text{ the } r\text{-th } \mathcal{C}^s\text{-conditional cumulant}) \end{aligned}$$

Special cases are

$$\begin{aligned} & \kappa_{\mathcal{C}^s,2}\Big(\sum_{k\in\mathcal{K}_s} u_k g_k(N_{s,k}^n/n_s)(N_{s,k}^n)^{-1/2} A_{j,k}\dot{\epsilon}_j + u_0 n_s^{-1/2}(\dot{\epsilon}_j^2 - \sigma_s^2)\Big) \\ = \; & E_{\mathcal{C}^s}\Big[\Big(\sum_{k\in\mathcal{K}_s} u_k g_k(N_{s,k}^n/n_s)(N_{s,k}^n)^{-1/2} A_{j,k}\dot{\epsilon}_j + u_0 n_s^{-1/2}(\dot{\epsilon}_j^2 - \sigma_s^2)\Big)^2\Big] \\ = \; & \sum_{k\in\mathcal{K}_t} u_k^2 g_k(N_{s,k}^n/n_s)^2 (N_{s,k}^n)^{-1} A_{j,k}\sigma_s^2 \quad (\kappa_{\mathcal{C}^s,2}[\dot{\epsilon}] = \sigma_s^2) \\ & + 2\sum_{k\in\mathcal{K}_s} u_k u_0 g_k(N_{s,k}^n/n_s)(N_{s,k}^n)^{-1/2} A_{j,k} n_s^{-1/2} E_{\mathcal{C}^s}[\dot{\epsilon}_1^3] \\ & + u_0^2 n_s^{-1}\big(E_{\mathcal{C}^s}[\dot{\epsilon}_1^4] - \sigma_s^4\big) \end{aligned}$$

and

$$
\begin{aligned}
& \kappa_{\mathcal{C}^s,3}\Bigg(\sum_{k\in\mathcal{K}_s} u_k g_k(N^n_{s,k}/n_s)(N^n_{s,k})^{-1/2}A_{j,k}\dot{\epsilon}_j + u_0 n_s^{-1/2}\sum_{j\in\mathbb{J}^n_s}(\dot{\epsilon}_j^2-\sigma_s^2)\Bigg) \\
= \; & E_{\mathcal{C}^s}\Bigg[\Bigg(\sum_{k\in\mathcal{K}_s} u_k g_k(N^n_{s,k}/n_s)(N^n_{s,k})^{-1/2}A_{j,k}\dot{\epsilon}_j + u_0 n_s^{-1/2}(\dot{\epsilon}_j^2-\sigma_s^2)\Bigg)^3\Bigg] \\
= \; & E_{\mathcal{C}^s}\Bigg[\Bigg(\sum_{k\in\mathcal{K}_s} u_k g_k(N^n_{s,k}/n_s)(N^n_{s,k})^{-1/2}A_{j,k}\dot{\epsilon}_j\Bigg)^3\Bigg] \\
& +3E_{\mathcal{C}^s}\Bigg[\Bigg(\sum_{k\in\mathcal{K}_s} u_k g_k(N^n_{s,k}/n_s)(N^n_{s,k})^{-1/2}A_{j,k}\dot{\epsilon}_j\Bigg)^2 u_0 n_s^{-1/2}(\dot{\epsilon}_j^2-\sigma_s^2)\Bigg] \\
& +3E_{\mathcal{C}^s}\Bigg[\Bigg(\sum_{k\in\mathcal{K}_s} u_k g_k(N^n_{s,k}/n_s)(N^n_{s,k})^{-1/2}A_{j,k}\dot{\epsilon}_j\Bigg)\Bigg(u_0 n_s^{-1/2}(\dot{\epsilon}_j^2-\sigma_s^2)\Bigg)^2\Bigg] \\
& +E_{\mathcal{C}^s}\Bigg[\Bigg(u_0 n_s^{-1/2}(\dot{\epsilon}_j^2-\sigma_s^2)\Bigg)^3\Bigg] \\
= \; & \sum_{k\in\mathcal{K}_s} u_k^3 g_k(N^n_{s,k}/n_s)^3(N^n_{s,k})^{-3/2}A_{j,k}E_{\mathcal{C}^s}[\dot{\epsilon}_j{}^3] \\
& +3\sum_{k\in\mathcal{K}_s} u_k^2 u_0 g_k(N^n_{s,k}/n_s)^2(N^n_{s,k})^{-1}n_s^{-1/2}A_{j,k}\big(E_{\mathcal{C}^s}\big[(\dot{\epsilon}_j)^4\big]-\sigma_s^4\big) \\
& +3\sum_{k\in\mathcal{K}_s} u_k u_0^2 g_k(N^n_{s,k}/n_s)(N^n_{s,k})^{-1/2}A_{j,k}n_s^{-1}\big(E_{\mathcal{C}^s}\big[\dot{\epsilon}_j{}^5\big]-2\sigma_s^2E_{\mathcal{C}^s}\big[\dot{\epsilon}_j{}^3\big]\big) \\
& +u_0^3 n_s^{-3/2}\big(E_{\mathcal{C}^s}\big[\dot{\epsilon}_j{}^6\big]-3\sigma_s^2E_{\mathcal{C}^s}\big[\dot{\epsilon}_j{}^4\big]+2\sigma_s^6\big)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\chi_{n_s,2,\mathcal{C}^s}(u) &= \sum_{j\in\mathbb{J}^n_s}\mathtt{i}^2\kappa_{\mathcal{C}^s,2}\Bigg(\sum_{k\in\mathcal{K}_s} u_k g_k(N^n_{s,k}/n_s)(N^n_{s,k})^{-1/2}A_{j,k}\dot{\epsilon}_j + u_0 n_s^{-1/2}(\dot{\epsilon}_j^2-\sigma_s^2)\Bigg) \\
&= \sum_{k\in\mathcal{K}_s}(\mathtt{i}u_k)^2 g_k(N^n_{s,k}/n_s)^2\sigma_s^2 \\
&\quad +2\sum_{k\in\mathcal{K}_s}(\mathtt{i}u_k)(\mathtt{i}u_0)g_k(N^n_{s,k}/n_s)(N^n_{s,k})^{1/2}n_s^{-1/2}E_{\mathcal{C}^s}[\dot{\epsilon}_1^3] \\
&\quad +(\mathtt{i}u_0)^2\big(E_{\mathcal{C}^s}[\dot{\epsilon}_1^4]-\sigma_s^4\big)
\end{aligned} \tag{87}
$$

and

$$
\begin{aligned}
\chi_{n_s,3,\mathcal{C}^s}(u) &= \sum_{j\in\mathbb{J}^n_s}\mathtt{i}^3\kappa_{\mathcal{C}^s,3}\Bigg(\sum_{k\in\mathcal{K}_s} u_k g_k(N^n_{s,k}/n_s)(N^n_{s,k})^{-1/2}A_{j,k}\dot{\epsilon}_j + u_0 n_s^{-1/2}(\dot{\epsilon}_j^2-\sigma_s^2)\Bigg) \\
&= \sum_{k\in\mathcal{K}_s}(\mathtt{i}u_k)^3 g_k(N^n_{s,k}/n_s)^3(N^n_{s,k})^{-1/2}E_{\mathcal{C}^s}\big[\dot{\epsilon}_1{}^3\big] \\
&\quad +3\sum_{k\in\mathcal{K}_s}(\mathtt{i}u_k)^2(\mathtt{i}u_0)g_k(N^n_{s,k}/n_s)^2n_s^{-1/2}\big(E_{\mathcal{C}^s}\big[\dot{\epsilon}_1{}^4\big]-\sigma_s^4\big) \\
&\quad +3\sum_{k\in\mathcal{K}_s}(\mathtt{i}u_k)(\mathtt{i}u_0)^2 g_k(N^n_{s,k}/n_s)(N^n_{s,k})^{1/2}n_s^{-1}\big(E_{\mathcal{C}^s}\big[\dot{\epsilon}_1{}^5\big]-2\sigma_s^2E_{\mathcal{C}^s}\big[\dot{\epsilon}_1{}^3\big]\big) \\
&\quad +(\mathtt{i}u_0)^3n_s^{-1/2}\big(E_{\mathcal{C}^s}\big[\dot{\epsilon}_1{}^6\big]-3\sigma_s^2E_{\mathcal{C}^s}\big[\dot{\epsilon}_1{}^4\big]+2\sigma_s^6\big).
\end{aligned} \tag{88}
$$

Let us consider normalized variable

$$
\mathfrak{T}_s^n = \frac{\dot{\mathbb{Z}}_s^n}{\widehat{\sigma}_s}
$$

for $s \in \mathbb{S}$, where $\widehat{\sigma}_s^2$ is a $\mathcal{G}_s^n$-measurable statistic. Suppose that $\widehat{\sigma}_s^2$ admits a stochastic expansion

$$\widehat{\sigma}_s^2 \quad = \quad n_s^{-1} \sum_{j \in \mathbb{J}_s^n} \dot{\epsilon}_j^2 + R_{n_s}^s$$

with $R_{n_s}^s = O_{L^{\infty-}}(n_s^{-1})$. Then $\mathfrak{T}_s^n$ has a stochastic expansion

$$\begin{aligned} \mathfrak{T}_s^n \quad &= \quad \sigma_s^{-1} \dot{\mathbb{Z}}_s^n - \frac{1}{2\sigma_s^3} \dot{\mathbb{Z}}_s^n \big(\widehat{\sigma}_s^2 - \sigma_s^2\big) + \cdots \\ &= \quad \sigma_s^{-1} \dot{\mathbb{Z}}_s^n - \frac{1}{2\sigma_s^3 n_s^{1/2}} \ddot{\mathbb{Z}}_s^n \dot{\mathbb{Z}}_s^n + \widetilde{R}_{n_s}^s, \end{aligned}$$

where $\widetilde{R}_{n_s}^s = O_{L^{\infty-}}(n_s^{-1})$. Remark that $\ddot{\mathbb{Z}}_s^n$ is a scalar random variable, while $\dot{\mathbb{Z}}_s^n$ is two-dimensional. We set

$$Y_s^n(z_s) \quad = \quad \sigma_s^{-1} \dot{z}_s - \frac{1}{2\sigma_s^3 n_s^{1/2}} \ddot{z}_s \dot{z}_s \tag{89}$$

for $z_s = (\dot{z}_s, \ddot{z}_s)$. The map Y_s^n is $\mathbb{R}^2$-valued. Then

$$\mathfrak{T}_s^n \quad = \quad Y_s^n(\mathbb{Z}_s^n) + \widetilde{R}_{n_s}^s. \tag{90}$$

If we are interested in a linear combination

$$\mathfrak{T}^n \quad = \quad \sum_{s \in \mathbb{S}} \dot{L}_s^n \mathfrak{T}_s^n \tag{91}$$

of $\mathfrak{T}_s^n$ with coefficient matrices $\dot{L}_s^n$, each of which is a function of the components of L_s^n, then computation of the expectation $E[g(\mathfrak{T}^n)]$ is reduced in that of $E\big[f(\underline{L}_S^n, \underline{\mathfrak{T}}_S^n)\big]$ for

$$f(\underline{l}_S, \underline{y}_S) \quad = \quad g\Big(\sum_{s \in \mathbb{S}} \dot{l}_s, y_s\Big),$$

where the argument $\dot{l}_s$ is for $\dot{L}_s^n$, and y_s for $\mathfrak{T}_s^n$. Moreover, $E\big[f(\underline{L}_S^n, \underline{\mathfrak{T}}_S^n)\big]$ can be approximated by $E\big[f(\underline{L}_S^n, \underline{\mathbb{Y}}_S^n)\big]$. Then the error caused by the shift $\widetilde{R}_{n_s}^s$ is included in the terms $\omega_2^{\Phi_{s,c}^{\cdot}}$ if g is measure theoretically continuous. Finally, the expectation $E\big[f(\underline{L}_S^n, \underline{\mathbb{Y}}_S^n)\big]$ can be approximated by using the backward scheme (75) incorporating the asymptotic expansion method.

C.0.3 A test statistic as $\mathfrak{T}^n$ of (91)

Zhang et al. [20]) studied a batched bandits model in which the reward (effect) R_j for the action A_j expressed by (1):

$$R_j \quad = \quad A_j^\star \beta_{s(j)} + \dot{\epsilon}_j \quad (j \in \mathbb{J}^n).$$

When the bandits have two arms (i.e., $\overline{k}_s = 2$), they proposed using the test statistic

$$\mathfrak{U}^n \quad = \quad S^{-1/2} \sum_{s \in \mathbb{S}} \frac{(N_{s,1}^n N_{s,2}^n)^{1/2}}{n_s^{1/2} \widehat{\sigma}_s} \big(\widehat{\Delta}_s - a\big)$$

for testing the null hypothesis that the treatment has no effect, i.e., H_0: $a = 0$, where the batched OLS $\widehat{\Delta}_s$ at batch s is defined by

$$\widehat{\Delta}_s \quad = \quad \frac{\sum_{j \in \mathbb{J}_s^n} A_{j,1} R_j}{\sum_{j \in \mathbb{J}_s^n} A_{j,1}} - \frac{\sum_{j \in \mathbb{J}_s^n} A_{j,2} R_j}{\sum_{j \in \mathbb{J}_s^n} A_{j,2}}.$$

Consider the functional $\mathfrak{T}_s^n = \dot{\mathbb{Z}}_s^n / \widehat{\sigma}_s$ with

$$\begin{aligned} \dot{\mathbb{Z}}_s^n \quad &= \quad \Big((N_{s,1}^n)^{-1/2} \sum_{j \in \mathbb{J}_s^n} A_{j,1} \dot{\epsilon}_j, (N_{s,2}^n)^{-1/2} \sum_{j \in \mathbb{J}_s^n} A_{j,2} \dot{\epsilon}_j \Big)^\star \\ &= \quad \begin{pmatrix} (N_{s,1}^n / n_s)^{-1/2} & 0 \\ 0 & (N_{s,2}^n / n_s)^{-1/2} \end{pmatrix} \begin{pmatrix} n_s^{-1/2} \sum_{j \in \mathbb{J}_s^n} A_{j,1} \dot{\epsilon}_j \\ n_s^{-1/2} \sum_{j \in \mathbb{J}_s^n} A_{j,2} \dot{\epsilon}_j \end{pmatrix} \\ &=: \quad G_{n_s}^s n_s^{-1/2} Z_{n_s}^s, \end{aligned}$$

$g_k(x) = 1$ for $k \in \mathcal{K}_s = \{1, 2\}$, and

$$\widehat{\sigma}_s^2 \quad =$$

Under H_0: $a = 0$, we have

$$\begin{aligned}
\mathfrak{T}^n &:= \mathfrak{U}^n \\
&= S^{-1/2} \sum_{s \in \mathbb{S}} \frac{(N_{s,1}^n N_{s,2}^n)^{1/2}}{n_s^{1/2} \widehat{\sigma}_s} \left(\frac{\sum_{j \in \mathbb{J}_s^n} A_{j,1} R_j}{\sum_{j \in \mathbb{J}_s^n} A_{j,1}} - \frac{\sum_{j \in \mathbb{J}_s^n} A_{j,2} R_j}{\sum_{j \in \mathbb{J}_s^n} A_{j,2}} \right) \\
&= S^{-1/2} \sum_{s \in \mathbb{S}} \frac{(N_{s,1}^n N_{s,2}^n)^{1/2}}{n_s^{1/2} \widehat{\sigma}_s} \left(\frac{\sum_{j \in \mathbb{J}_s^n} A_{j,1} \dot{\epsilon}_j}{\sum_{j \in \mathbb{J}_s^n} A_{j,1}} - \frac{\sum_{j \in \mathbb{J}_s^n} A_{j,2} \dot{\epsilon}_j}{\sum_{j \in \mathbb{J}_s^n} A_{j,2}} \right) \\
&= S^{-1/2} \sum_{s \in \mathbb{S}} \frac{(N_{s,1}^n N_{s,2}^n)^{1/2}}{n_s^{1/2} \widehat{\sigma}_s} \left(\frac{(N_{s,1}^n)^{1/2}}{\sum_{j \in \mathbb{J}_s^n} A_{j,1}} (N_{s,1}^n)^{-1/2} \sum_{j \in \mathbb{J}_s^n} A_{j,1} \dot{\epsilon}_j \right. \\
&\qquad \left. - \frac{(N_{s,22}^n)^{1/2}}{\sum_{j \in \mathbb{J}_s^n} A_{j,2}} (N_{s,2}^n)^{-1/2} \sum_{j \in \mathbb{J}_s^n} A_{j,2} \dot{\epsilon}_j \right) \\
&= S^{-1/2} \sum_{s \in \mathbb{S}} \frac{(N_{s,1}^n N_{s,2}^n)^{1/2}}{n_s^{1/2}} \left((N_{s,1}^n)^{-1/2}, \, -(N_{s,2}^n)^{-1/2} \right) \mathfrak{T}_s^n \\
&= \sum_{s \in \mathbb{S}} \dot{L}_s^n \mathfrak{T}_s^n
\end{aligned}$$

for

$$\dot{L}_s^n \quad = \quad S^{-1/2} \big((N_{s,2}^n / n_s)^{1/2}, \, -(N_{s,1}^n / n_s)^{1/2} \big). \tag{92}$$

Therefore, we can apply the asymptotic expansion for $\mathfrak{T}^n$ of (91) to $\mathfrak{U}^n$.

Suppose that

$$E_{\mathcal{C}^s}[\dot{\epsilon}_j] \; = \; 0, \quad E_{\mathcal{C}^s}[\dot{\epsilon}_j^2] \; = \; \sigma_s^2, \quad E_{\mathcal{C}^s}[\dot{\epsilon}_j^r] \; = \; \mu_{s,r} \quad (r = 3, ..., 6)$$

for $j \in \mathbb{J}_s^\infty$, for some constants σ_s^2 and $\mu_{s,r}$. From (87) and (88) for $g_k(x) = 1$, we obtain

$$\begin{aligned}
\chi_{n_s,2,\mathcal{C}^s}(u) \quad = \quad & \sum_{k \in \mathcal{K}_s} (\mathtt{i} u_k)^2 \sigma_s^2 + 2 \sum_{k \in \mathcal{K}_s} (\mathtt{i} u_k)(\mathtt{i} u_0)(N_{s,k}^n)^{1/2} n_s^{-1/2} \mu_{s,3} \\
& + (\mathtt{i} u_0)^2 \big(\mu_{s,4} - \sigma_s^4 \big)
\end{aligned} \tag{93}$$

and

$$\begin{aligned}
\chi_{n_s,3,\mathcal{C}^s}(u) \quad = \quad & \sum_{k \in \mathcal{K}_s} (\mathtt{i} u_k)^3 (N_{s,k}^n)^{-1/2} \mu_{s,3} \\
& + 3 \sum_{k \in \mathcal{K}_s} (\mathtt{i} u_k)^2 (\mathtt{i} u_0) n_s^{-1/2} \big(\mu_{s,4} - \sigma_s^4 \big) \\
& + 3 \sum_{k \in \mathcal{K}_s} (\mathtt{i} u_k)(\mathtt{i} u_0)^2 (N_{s,k}^n)^{1/2} n_s^{-1} \big(\mu_{s,5} - 2\sigma_s^2 \mu_{s,3} \big) \\
& + u_0^3 n_s^{-1/2} \big(\mu_{s,6} - 3\sigma_s^2 \mu_{s,4} + 2\sigma_s^6 \big).
\end{aligned} \tag{94}$$

For the random matrix

$$V_s^{n_s} \quad = \quad \begin{pmatrix} \sigma_s^2 & 0 & (N_{s,1}^n)^{1/2} n_s^{-1/2} \mu_{s,3} \\ 0 & \sigma_s^2 & (N_{s,2}^n)^{1/2} n_s^{-1/2} \mu_{s,3} \\ (N_{s,1}^n)^{1/2} n_s^{-1/2} \mu_{s,3} & (N_{s,2}^n)^{1/2} n_s^{-1/2} \mu_{s,3} & \mu_{s,4} - \sigma_s^4 \end{pmatrix},$$

the Hermite polynomials $H_{\alpha_1,\alpha_2,\alpha_3}$ are defined by

$$\begin{aligned}
H_{\alpha_1,\alpha_2,\alpha_3}(z_1, z_2, z_0; V_s^{n_s}) \quad = \quad & \phi\big((z_1, z_2, z_0); 0, V_s^{n_s}\big)^{-1} \\
& \times (-\partial_{z_{\alpha_1}})(-\partial_{z_{\alpha_2}})(-\partial_{z_{\alpha_3}}) \phi\big((z_1, z_2, z_0); 0, V_s^{n_s}\big)
\end{aligned}$$

for $(\alpha_1, \alpha_2, \alpha_3) \in \{1, 2, 0\}^3$.

Now, the first-order asymptotic expansion formula for $\mathcal{L}\{\mathbb{Z}_s^n \,|\mathcal{C}^s\}$ is obtained from (69) as

$$\begin{aligned}
\frac{d\Psi^n_{s,\mathsf{p},\mathbf{w}^n_s}}{dz}(z) &= \phi\big((z_1, z_2, z_0); 0, V_s^{n_s}\big) \\
&\quad + \frac{1}{6} n_s^{-1/2} \phi\big((z_1, z_2, z_0); 0, V_s^{n_s}\big) \\
&\quad \times \Bigg\{ \sum_{k \in \mathcal{K}_s} (N^n_{s,k})^{-1/2} n_s^{1/2} \mu_{s,3} H_{k,k,k}(z_1, z_2, z_0; V_s^{n_s}) \\
&\qquad + 3 \sum_{k \in \mathcal{K}_s} \big(\mu_{s,4} - \sigma_s^4\big) H_{k,k,0}(z_1, z_2, z_0; V_s^{n_s}) \\
&\qquad + 3 \sum_{k \in \mathcal{K}_s} (N^n_{s,k})^{1/2} n_s^{-1/2} \big(\mu_{s,5} - 2\sigma_s^2 \mu_{s,3}\big) H_{k,0,0}(z_1, z_2, z_0; V_s^{n_s}) \\
&\qquad + \big(\mu_{s,6} - 3\sigma_s^2 \mu_{s,4} + 2\sigma_s^6\big) H_{0,0,0}(z_1, z_2, z_0; V_s^{n_s}) \Bigg\} \qquad (95)
\end{aligned}$$

for $z = (z_1, z_2, z_0) \in \mathbb{R}^3$.

For $\dot{L}_s^n$ of (92), let us consider the probability

$$\begin{aligned}
P\big[\mathfrak{T}^n \geq x\big] &= P\Bigg[\sum_{s \in \mathbb{S}} \dot{L}_s^n \mathfrak{T}_s^n \geq x\Bigg] \qquad ((91)) \\
&= P\Bigg[\sum_{s \in \mathbb{S}} \dot{L}_s^n Y_s^n(\mathbb{Z}_s^n) + \widetilde{R}^s_{n_s} \geq x\Bigg] \qquad ((90)) \\
&= P\Bigg[\sum_{s \in \mathbb{S}} \dot{L}_s^n Y_s^n(\mathbb{Z}_s^n) \geq x\Bigg] + \overline{O}(n^{-1})
\end{aligned}$$

where $\overline{O}(n^{-1})$ denotes a term that is of order $O(n^{-1})$ uniformly in $x \in \mathbb{R}$.

In this situation, f of (21) is

$$f\big(\underline{l}_S, \underline{y}_S\big) = 1_{\big\{\sum_{s \in \mathbb{S}} \dot{l}_s y_s \geq x\big\}} \qquad \big(\underline{l}_S = \big(l_s = \big(\dot{l}_s, \ddot{l}_s\big)\big)_{s \in \mathbb{S}},\ \underline{y}_S = (y_s)_{s \in \mathbb{S}}\big)$$

and set $\overline{L}_S^n = \big(\dot{L}_s^n\big)_{s \in \mathbb{S}}$ with $\dot{L}_s^n$ of (92), as well as $\underline{\mathbb{Y}}_S^n = \big(\mathbb{Y}_s^n\big)_{s \in \mathbb{S}}$ with

$$\mathbb{Y}_s^n = Y_s^n(\mathbb{Z}_s^n)$$

for the function Y_s^n of (89) and $\mathbb{Z}_s^n$ of (84) accompanied with (**??**) and (85).

Let us illustrate the backward approximation in the case $S = 2$ though the description is quite the same for a larger value of S.

$$\widehat{f}^n_{2,c}(\underline{l}_2, \underline{z}_2) = 1_{\big\{\sum_{s=1}^2 \dot{l}_s Y_s^n(z_s) \geq x\big\}}$$

From (75),

$$
\begin{aligned}
& \widehat{f}_1^n(\underline{l}_1, \underline{z}_1) \\
= & \int \widehat{f}_2^n\big(\underline{l}_2, \underline{z}_2\big) \Psi^n_{2,\mathsf{p},\mathbf{w}_2^n}(dz_2) 1_{\{l_2 \in \mathcal{A}_2^n\}} \nu^n_{c_2}(dl_2, d\mathbf{w}_2^n) q\big((\underline{l}_1, \underline{z}_1), dc_2\big) \\
= & \int 1_{\left\{ \sum_{s=1}^2 \dot{l}_s Y_s^n(z_s) \geq x \right\}} \Psi^n_{2,\mathsf{p},\mathbf{w}_2^n}(dz_2) 1_{\{l_2 \in \mathcal{A}_2^n\}} \nu^n_{c_2}(dl_2, d\mathbf{w}_2^n) q\big((\underline{l}_1, \underline{z}_1), dc_2\big) \\
= & \int \int_{\substack{\dot{z}_2 \in \mathbb{R}^2 \\ \ddot{z}_2 \in \mathbb{R}}} 1_{\left\{ \sum_{s=1}^2 \dot{l}_s \left(\sigma_s^{-1} \dot{z}_s - \frac{1}{2\sigma_s^3 n_s^{1/2}} \ddot{z}_s \dot{z}_s \right) \geq x \right\}} \Psi^n_{2,\mathsf{p},\mathbf{w}_2^n}(d\dot{z}_2, d\ddot{z}_2) \\
& \qquad \times 1_{\{l_2 \in \mathcal{A}_2^n\}} \nu^n_{c_2}(dl_2, d\mathbf{w}_2^n) q\big((\underline{l}_1, \underline{z}_1), dc_2\big) \\
= & \int \int_{\substack{\dot{z}_2 \in \mathbb{R}^2 \\ \ddot{z}_2 \in \mathbb{R}}} 1_{\left\{ \dot{l}_2 \left(\sigma_2^{-1} \dot{z}_2 - \frac{1}{2\sigma_2^3 n_2^{1/2}} \ddot{z}_2 \dot{z}_2 \right) \geq x - \dot{l}_1 \left(\sigma_1^{-1} \dot{z}_1 - \frac{1}{2\sigma_1^3 n_1^{1/2}} \ddot{z}_1 \dot{z}_1 \right) \right\}} \Psi^n_{2,\mathsf{p},\mathbf{w}_2^n}(d\dot{z}_2, d\ddot{z}_2) \\
& \qquad \times 1_{\{l_2 \in \mathcal{A}_{2,c}^n\}} \nu^n_{c_2}(dl_2, d\mathbf{w}_2^n) q\big((\underline{l}_1, \underline{z}_1), dc_2\big)
\end{aligned}
$$

(Only at this batch 2, we can apply the transform of the density if you like.)

$$
\begin{aligned}
\widehat{f}_0^n &= \int \widehat{f}_1^n\big(\underline{l}_1, \underline{z}_1\big) 1_{\left\{ (\underline{l}_1, \underline{z}_1) \in B(c_2) \right\}} \Psi^n_{1,\mathsf{p},\mathbf{w}_1^n}(dz_1) 1_{\{l_1 \in \mathcal{A}_{1,c}^n\}} \nu^n_{c_1}(dl_1, d\mathbf{w}_1^n) \\
&= \int \int_{\substack{\dot{z}_1 \in \mathbb{R}^2 \\ \ddot{z}_1 \in \mathbb{R}}} \int \int_{\substack{\dot{z}_2 \in \mathbb{R}^2 \\ \ddot{z}_2 \in \mathbb{R}}} 1_{\left\{ \sum_{s=1}^2 \dot{l}_s \left(\sigma_s^{-1} \dot{z}_s - \frac{1}{2\sigma_s^3 n_s^{1/2}} \ddot{z}_s \dot{z}_s \right) \geq x \right\}} \Psi^n_{2,\mathsf{p},\mathbf{w}_2^n}(d\dot{z}_2, d\ddot{z}_2) \\
&\qquad \times 1_{\{l_2 \in \mathcal{A}_{2,c}^n\}} \nu^n_{c_2}(dl_2, d\mathbf{w}_2^n) q\big((\underline{l}_1, \underline{z}_1), dc_2\big) \\
&\qquad \times \Psi^n_{1,\mathsf{p},\mathbf{w}_1^n}(d\dot{z}_1, d\ddot{z}_1) 1_{\{l_1 \in \mathcal{A}_{1,c}^n\}} \nu^n_{c_1}(dl_1, d\mathbf{w}_1^n)
\end{aligned}
$$

By Fubini's theorem,

$$
\begin{aligned}
\widehat{f}_0^n &= \int \int \int_{\substack{\dot{z}_1 \in \mathbb{R}^2 \\ \ddot{z}_1 \in \mathbb{R}}} \int_{\substack{\dot{z}_2 \in \mathbb{R}^2 \\ \ddot{z}_2 \in \mathbb{R}}} \int_{c_2 \in \mathfrak{C}_2} 1_{\left\{ \sum_{s=1}^2 \dot{l}_s \left(\sigma_s^{-1} \dot{z}_s - \frac{1}{2\sigma_s^3 n_s^{1/2}} \ddot{z}_s \dot{z}_s \right) \geq x \right\}} q\big((l_1, z_1), dc_2\big) \\
&\qquad \times \Psi^n_{2,\mathsf{p},\mathbf{w}_2^n}(d\dot{z}_2, d\ddot{z}_2) \Psi^n_{1,\mathsf{p},\mathbf{w}_2^{n_1}}(d\dot{z}_1, d\ddot{z}_1) \\
&\qquad \times 1_{\{l_2 \in \mathcal{A}_{2,c}^n\}} \nu^n_{c_2}(dl_2, d\mathbf{w}_2^n) 1_{\{l_1 \in \mathcal{A}_{1,c}^n\}} \nu^n_{c_1}(dl_1, d\mathbf{w}_1^n)
\end{aligned}
$$

An example is the ϵ-Greedy algorithm having $\mathfrak{C}_2 = \{c_2^{(1)}, c_2^{(2)}\}$ with

$$
q\big((\underline{l}_1, \underline{z}_1), dc_2\big) = 1_{\{\mathtt{h}_1(\underline{l}_1, \underline{z}_1) \geq 0\}} \delta_{c_2^{(1)}}(dc_2) + 1_{\{\mathtt{h}_1(\underline{l}_1, \underline{z}_1) < 0\}} \delta_{c_2^{(2)}}(dc_2)
$$

for some function $\mathtt{h}_1$. On the other hand, a Thompson sampling is realized as

$$
\begin{aligned}
q\big((\underline{l}_1, \underline{z}_1), dc_2\big) &= 1_{\{\mathtt{h}_1(\underline{l}_1, \underline{z}_1) \leq a_1\}} \delta_{c_2^{(1)}}(dc_2) + 1_{\{\mathtt{h}_1(\underline{l}_1, \underline{z}_1) > a_2\}} \delta_{c_2^{(2)}}(dc_2) \\
&\quad + 1_{\{a_1 < \mathtt{h}_1(\underline{l}_1, \underline{z}_1) \leq a_2\}} \delta_{C_2(l_1, z_1)}(dc_2)
\end{aligned} \tag{96}
$$

with some constants a_1, a_2 and some $\mathfrak{C}_2$-valued function C_2 of (l_1, z_1). Obviously, the right-hand side of (96) can be written as $\delta_{C_2'(l_1, z_1)}(dc_2)$ with a $\mathfrak{C}_2$-valued function C_2' of (l_1, z_1) defined by

$$
C_2'(l_1, z_1) = \begin{cases} c_2^{(1)} & (\mathtt{h}_1(l_1, z_1) \leq a_1) \\ C_2(l_1, z_1) & (a_1 < \mathtt{h}_1(\underline{l}_1, \underline{z}_1) \leq a_2) \\ c_2^{(2)} & (\mathtt{h}_1(\underline{l}_1, \underline{z}_1) > a_2). \end{cases}
$$

In the ϵ-greedy algorithm, the strategy is determined by the sign of the statistic

$$
\Xi_1^n = (N_{1,1}^n)^{-1} \sum_{j \in \mathbb{J}_1^n} A_{j,1} R_j - (N_{1,2}^n)^{-1} \sum_{j \in \mathbb{J}_1^n} A_{j,2} R_j
$$

Under the null hypothesis H_0: $a = 0$,

$$\begin{aligned} n^{1/2}\Xi_1^n &= n^{1/2}(N_{1,1}^n)^{-1}\sum_{j\in\mathbb{J}_1^n} A_{j,1}\dot{\epsilon}_j - n^{1/2}(N_{1,2}^n)^{-1}\sum_{j\in\mathbb{J}_1^n} A_{j,2}\dot{\epsilon}_j \\ &= \ddot{L}_1^n \cdot \dot{\mathbb{Z}}_1^n \end{aligned}$$

where

$$\ddot{L}_1^n = \left((N_{1,1}^n/n)^{-1/2}, -(N_{1,2}^n/n)^{-1/2}\right)$$

Set

$$\begin{aligned} L_1^n &= (\dot{L}_1^n, \ddot{L}_1^n), \\ L_2^n &= (\dot{L}_2^n, \ddot{L}_2^n) = (\dot{L}_2^n, 0), \\ l_1 &= (\dot{l}_1, \ddot{l}_1), \end{aligned}$$

Then

$$\begin{aligned} \widehat{f}_0^n = \int\int\int_{\substack{\dot{z}_1\in\mathbb{R}^2\\ \ddot{z}_1\in\mathbb{R}}}\int_{\substack{\dot{z}_2\in\mathbb{R}^2\\ \ddot{z}_2\in\mathbb{R}}}\int_{c_2\in\mathfrak{C}_2} & 1_{\left\{\sum_{s=1}^2 \dot{l}_s\left(\sigma_s^{-1}\dot{z}_s - \frac{1}{2\sigma_s^3 n_s^{1/2}}\ddot{z}_s\dot{z}_s\right)\geq x\right\}} q\big((l_1,z_1),dc_2\big) \\ & \times\Psi_{2,\mathsf{p},\mathbf{w}_2^n}^n(d\dot{z}_2,d\ddot{z}_2)\Psi_{1,\mathsf{p},\mathbf{w}_1^n}^n(d\dot{z}_1,d\ddot{z}_1) \\ & \times 1_{\{l_2\in\mathcal{A}_{2,c}^n\}}\nu_{c_2}^n(dl_2,d\mathbf{w}_2^n)1_{\{l_1\in\mathcal{A}_{1,c}^n\}}\nu_{c_1}^n(dl_1,d\mathbf{w}_1^n) \end{aligned} \tag{97}$$

For example, if $\mathtt{h}_1(l_1,z_1) = \ddot{l}_1\dot{z}_1$ in the ϵ-Greedy, then one should compute the components

$$\begin{aligned} \widehat{f}_{0,c}^n = \int\int\int_{\substack{\dot{z}_1\in\mathbb{R}^2\\ \ddot{z}_1\in\mathbb{R}}}\int_{\substack{\dot{z}_2\in\mathbb{R}^2\\ \ddot{z}_2\in\mathbb{R}}} & 1_{\left\{\sum_{s=1}^2 \dot{l}_s\left(\sigma_s^{-1}\dot{z}_s - \frac{1}{2\sigma_s^3 n_s^{1/2}}\ddot{z}_s\dot{z}_s\right)\geq x\right\}} 1_{\left\{(-1)^{c_2+1}\ddot{l}_1\dot{z}_1>0\right\}} \\ & \times\Psi_{2,\mathsf{p},\mathbf{w}_2^n}^n(d\dot{z}_2,d\ddot{z}_2)\Psi_{1,\mathsf{p},\mathbf{w}_1^n}^n(d\dot{z}_1,d\ddot{z}_1) \\ & \times 1_{\{l_2\in\mathcal{A}_{2,c}^n\}}\nu_{c_2}^n(dl_2,d\mathbf{w}_2^n)1_{\{l_1\in\mathcal{A}_{1,c}^n\}}\nu_{c_1}^n(dl_1,d\mathbf{w}_1^n) \end{aligned} \tag{98}$$

for $c = (c_1, c_2)$ with $c_2 \in \mathfrak{C}_2$ and a fixed $c_1 \in \mathfrak{C}_1$.

Introduce the transform

$$y_s = \begin{pmatrix}\dot{y}_s\\ \ddot{y}_s\end{pmatrix} = \begin{pmatrix}\sigma_s^{-1}\dot{z}_s\\ \sigma_s^{-2}\ddot{z}_s\end{pmatrix}. \tag{99}$$

Then the formula (97) has the expression

$$\begin{aligned} \widehat{f}_0^n = \int\int\int_{\substack{\dot{y}_1\in\mathbb{R}^2\\ \ddot{y}_1\in\mathbb{R}}}\int_{\substack{\dot{y}_2\in\mathbb{R}^2\\ \ddot{y}_2\in\mathbb{R}}}\int_{\mathfrak{C}_2} & 1_{\left\{\sum_{s=1}^2 \dot{l}_s\left(\dot{y}_s - \frac{1}{2n_s^{1/2}}\ddot{y}_s\dot{y}_s\right)\geq x\right\}} \\ & \times\widetilde{\Psi}_{2,\mathsf{p},\mathbf{w}_2^n}^n(d\dot{y}_2,d\ddot{y}_2)1_{\{l_2\in\mathcal{A}_{2,c}^n\}}\nu_{c_2}^n(dl_2,d\mathbf{w}_2^n) \\ & \times q\big((l_1,\sigma_1^{-1}y_1),dc_2\big)\widetilde{\Psi}_{1,\mathsf{p},\mathbf{w}_1^n}^n(d\dot{y}_1,d\ddot{y}_1)1_{\{l_1\in\mathcal{A}_{1,c}^n\}}\nu_{c_1}^n(dl_1,d\mathbf{w}_1^n) \end{aligned} \tag{100}$$

where $\widetilde{\Psi}_{s,\mathsf{p},\mathbf{w}_s^n}^n(d\dot{y}_s,d\ddot{y}_s)$ is the asymptotic expansion transformed by (99) from $\Psi_{s,\mathsf{p},\mathbf{w}_s^n}^n(d\dot{z}_s,d\ddot{z}_s)$. It should be remarked that each integral with respect to $\nu_{c_s}^n(dl_s,d\mathbf{w}_s^n)$ $(s=1,2)$ can be simplified to a more concise integral by using an aggregate statistic appearing in $\widetilde{\Psi}_{s,\mathsf{p},\mathbf{w}_s^n}^n(d\dot{y}_s,d\ddot{y}_s)$ than $\mathbf{w}^{n_s}$.

C.0.4 Transform of the asymptotic expansion

We will consider an analytic method for numerical computation of $\widehat{f}_0^n$. From (100),

$$\begin{aligned} \widehat{f}_0^n = \int\int\int_{\substack{\dot{y}_1\in\mathbb{R}^2\\ \ddot{y}_1\in\mathbb{R}}}\int_{\mathfrak{C}_2} & \mathcal{I}_2(\dot{y}_1,\ddot{y}_1)1_{\{l_2\in\mathcal{A}_{2,c}^n\}}\nu_{c_2}^n(dl_2,d\mathbf{w}_2^n) \\ & \times q\big((l_1,\sigma_1^{-1}y_1),dc_2\big)\psi_1(\mathbf{w}_1^n,\dot{y}_1,\ddot{y}_1)d\dot{y}_1d\ddot{y}_1 1_{\{l_1\in\mathcal{A}_{1,c}^n\}}\nu_{c_1}^n(dl_1,d\mathbf{w}_1^n), \end{aligned} \tag{101}$$

where

$$\mathcal{I}_2(\dot{y}_1,\ddot{y}_1) \quad = \quad \int_{\substack{\dot{y}_2\in\mathbb{R}^2\\ \ddot{y}_2\in\mathbb{R}}} 1_{\left\{\sum_{s=1}^2 \dot{l}_s\left(\dot{y}_s-\frac{1}{2n_s^{1/2}}\ddot{y}_s\dot{y}_s\right)\geq x\right\}}\psi_2(\mathbf{w}_2^n,\dot{y}_2,\ddot{y}_2)d\dot{y}_2d\ddot{y}_2$$

for given $\underline{l}_2$ and $\underline{\mathbf{w}}_2^n$, where

$$\begin{aligned}\psi_1(\mathbf{w}_1^n,\dot{y}_1,\ddot{y}_1) &= \frac{d\widetilde{\Psi}^n_{1,\mathsf{p},\mathbf{w}_1^n}}{dy_1}(\dot{y}_1,\ddot{y}_1)\\ \psi_2(\mathbf{w}_2^n,\dot{y}_2,\ddot{y}_2) &= \frac{d\widetilde{\Psi}^n_{2,\mathsf{p},\mathbf{w}_2^n}}{dy_2}(\dot{y}_2,\ddot{y}_2)\end{aligned}$$

for $\underline{l}_2$ and $\underline{\mathbf{w}}_2^n$. The formula (101) is rewritten as

$$\mathcal{I} \quad = \quad \int_{\substack{\dot{y}_1\in\mathbb{R}^2\\ \ddot{y}_1\in\mathbb{R}}} \mathcal{I}_2(\dot{y}_1,\ddot{y}_1)\widetilde{q}\big((l_1,\sigma_1^{-1}y_1),c_2\big)\psi_1(\dot{y}_1,\ddot{y}_1)d\dot{y}_1d\ddot{y}_1 \tag{102}$$

with

$$\mathcal{I}_2(\dot{y}_1,\ddot{y}_1) \quad = \quad \int_{\substack{\dot{y}_2\in\mathbb{R}^2\\ \ddot{y}_2\in\mathbb{R}}} 1_{\left\{\sum_{s=1}^2 \dot{l}_s\left(\dot{y}_s-\frac{1}{2n_s^{1/2}}\ddot{y}_s\dot{y}_s\right)\geq x\right\}}\psi_2(\dot{y}_2,\ddot{y}_2)d\dot{y}_2d\ddot{y}_2.$$

Let $x>0$. For numerical computation of $\widehat{f}^n_{0,c}$, we need to compute the value of the integral

$$\begin{aligned}\mathcal{I} \quad &= \quad \int_{\substack{\dot{y}_1\in\mathbb{R}^2\\ \ddot{y}_1\in\mathbb{R}}}\int_{\substack{\dot{y}_2\in\mathbb{R}^2\\ \ddot{y}_2\in\mathbb{R}}} 1_{\left\{\sum_{s=1}^2 \dot{l}_s\left(\dot{y}_s-\frac{1}{2n_s^{1/2}}\ddot{y}_s\dot{y}_s\right)\geq x\right\}}1_{\left\{(-1)^{c_2+1}\ddot{l}_1\dot{y}_1>0\right\}}\\ &\qquad\qquad\times\psi_2(\dot{y}_2,\ddot{y}_2)d\dot{y}_2d\ddot{y}_2\psi_1(\dot{y}_1,\ddot{y}_1)d\dot{y}_1d\ddot{y}_1\\ &= \quad \int_{\substack{\dot{y}_1\in\mathbb{R}^2\\ \ddot{y}_1\in\mathbb{R}}} \mathcal{I}_2(\dot{y}_1,\ddot{y}_1)1_{\left\{(-1)^{c_2+1}\ddot{l}_1\dot{y}_1>0\right\}}\psi_1(\dot{y}_1,\ddot{y}_1)d\dot{y}_1d\ddot{y}_1\end{aligned} \tag{103}$$

for given c, $\underline{l}_2$ and $\underline{\mathbf{w}}_2^n$, where

$$\begin{aligned}\psi_1(\dot{y}_1,\ddot{y}_1) &= \frac{d\widetilde{\Psi}^n_{1,\mathsf{p},\mathbf{w}_1^n}}{dy_1}(\dot{y}_1,\ddot{y}_1)\\ \psi_2(\dot{y}_2,\ddot{y}_2) &= \frac{d\widetilde{\Psi}^n_{2,\mathsf{p},\mathbf{w}_2^n}}{dy_2}(\dot{y}_2,\ddot{y}_2)\end{aligned}$$

for $\underline{l}_2$ and $\underline{\mathbf{w}}_2^n$, and

$$\mathcal{I}_2(\dot{y}_1,\ddot{y}_1) \quad = \quad \int_{\substack{\dot{y}_2\in\mathbb{R}^2\\ \ddot{y}_2\in\mathbb{R}}} 1_{\left\{\sum_{s=1}^2 \dot{l}_s\left(\dot{y}_s-\frac{1}{2n_s^{1/2}}\ddot{y}_s\dot{y}_s\right)\geq x\right\}}\psi_2(\dot{y}_2,\ddot{y}_2)d\dot{y}_2d\ddot{y}_2.$$

The integral $\mathcal{I}_2(\dot{y}_1,\ddot{y}_1)$ is rewritten as

$$\mathcal{I}_2(\dot{y}_1,\ddot{y}_1) \quad = \quad \int_{\substack{\dot{y}_2\in\mathbb{R}^2\\ \ddot{y}_2\in\mathbb{R}}} 1_{\left\{\dot{l}_2\left(\dot{y}_2-\frac{1}{2n_2^{1/2}}\ddot{y}_2\dot{y}_2\right)\geq U_1(\dot{y}_1,\ddot{y}_1)\right\}}\psi_2(\dot{y}_2,\ddot{y}_2)d\dot{y}_2d\ddot{y}_2 \tag{104}$$

with

$$U_1(\dot{y}_1,\ddot{y}_1) \quad = \quad x-\dot{l}_1\big(\dot{y}_1-\frac{1}{2n_1^{1/2}}\ddot{y}_1\dot{y}_1\big).$$

We change variables by

$$\begin{pmatrix}\dot{w}\\ \ddot{w}\end{pmatrix} \quad = \quad F(\dot{y}_2,\ddot{y}_2) \quad = \quad \begin{pmatrix}\widetilde{l}_2\big(\dot{y}_2-\frac{1}{2n_2^{1/2}}\ddot{y}_2\dot{y}_2\big)\\ \ddot{y}_2\end{pmatrix}, \tag{105}$$

where

$$\widetilde{l}_2 \;=\; \begin{pmatrix} \dot{l}_2 \\ \dot{l}_2^{\perp} \end{pmatrix} \tag{106}$$

with

$$\dot{l}_2^{\perp} \;=\; \big((N_{2,1}^n/n_2)^{1/2}, (N_{2,2}^n/n_2)^{1/2}\big)$$

for example. Write $\dot{w} = (\dot{w}^{(1)}, \dot{w}^{(2)})$. The change of variables (105) applied to (104) yields

$$\begin{aligned}
\mathcal{I}_2(\dot{y}_1, \ddot{y}_1) \;&=\; \int_{\substack{\dot{y}_2 \in \mathbb{R}^2 \\ \ddot{y}_2 \in \mathbb{R}}} 1_{\left\{ \dot{l}_2 \left(\dot{y}_2 - \frac{1}{2n_2^{1/2}} \ddot{y}_2 \dot{y}_2 \right) \geq U_1(\dot{y}_1, \ddot{y}_1) \right\}} \psi_2(\dot{y}_2, \ddot{y}_2) d\dot{y}_2 d\ddot{y}_2 \qquad (107) \\
&=\; \int_{\substack{\dot{y}_2 \in \mathbb{R}^2 \\ \ddot{y}_2 \in \mathbb{R}}} 1_{\left\{ \dot{w}^{(1)} \geq U_1(\dot{y}_1, \ddot{y}_1) \right\}} \psi_2(F^{-1}(\dot{w}, \ddot{w})) \big| \det J_{F^{-1}}(\dot{w}, \ddot{w}) \big| d\dot{w} d\ddot{w} \\
&=\; \int_{\substack{\dot{y}_2 \in \mathbb{R}^2 \\ \ddot{y}_2 \in \mathbb{R}}} 1_{\left\{ \dot{w}^{(1)} \geq U_1(\dot{y}_1, \ddot{y}_1) \right\}} \psi_2\big(\tilde{l}_2^{-1} \dot{w} + 2^{-1} n_2^{-1/2} \ddot{w} \tilde{l}_2^{-1} \dot{w}, \ddot{w}\big) \\
&\qquad\qquad \times \big| \det \big((1 + 2^{-1} n_2^{-1/2} \ddot{w}) \widetilde{l}_2^{-1} \big) \big| d\dot{w} d\ddot{w} + o(n_s^{-1/2-\delta}) \\
&=:\; \widetilde{\mathcal{I}}_2(\dot{y}_1, \ddot{y}_1) + o(n_s^{-1/2-\delta}) \qquad (108)
\end{aligned}$$

as $n_s \to \infty$ for some positive constant δ. Here $J_{F^{-1}}$ is the Jacobian of the map F^{-1}. The passage from (107) to (108) is intuitive because the map F^{-1} is not well defined on the whole space. However, the approximation $\widetilde{\mathcal{I}}_2(\dot{y}_1, \ddot{y}_1)$ at (108) to $\mathcal{I}_2(\dot{y}_1, \ddot{y}_1)$ is valid and the order of the error is uniform in $(\dot{y}_1, \ddot{y}_1)$. More precisely, first we restrict the domain of the integral to $\big\{(\dot{y}_2, \ddot{y}_2);\ |(\dot{y}_2, \ddot{y}_2)| < n_2^{1/4}\big\}$, then the map F becomes a diffeomorphism and the above computation is validated for large n_2. The contribution of the integral over the cut off area is negligible. As a matter of fact, this approach was taken by Bhattacharya and Ghosh [3], cf. Sakamoto and Yoshida [14, 15], Yoshida [19]. The approximation formula $\widetilde{\mathcal{I}}_2(\dot{y}_1, \ddot{y}_1)$ at (108) can further be simplified by expanding ψ_2 and $\det$, and neglecting the terms of order n_2^{-1} or smaller. It should be remarked that thus simplified $\widetilde{\mathcal{I}}_2(\dot{y}_1, \ddot{y}_1)$ is a smooth function of $(\dot{y}_1, \ddot{y}_1)$ and has a closed form.

For the computation of the integral of $\mathcal{I}_2(\dot{y}_1, \ddot{y}_1)$ of (104), we may apply a linear transform of $(\dot{y}_1, \ddot{y}_1)$ with a matrix $\widetilde{l}_1$ like (106). Since the function $\mathcal{I}_2(\dot{y}_1, \ddot{y}_1)$ is in $C_b^\infty(\mathbb{R}^3)$, an approximation by a power series in $(\dot{y}_1, \ddot{y}_1)$ should work well. Then the integral of $\mathcal{I}_2(\dot{y}_1, \ddot{y}_1)$ can be calculated with an closed form of approximation. Our approach is feasible for S larger than two.

C.0.5 Importance sampling

Let

$$\begin{aligned}
\mathcal{J}(\dot{y}_1, \ddot{y}_1, \dot{y}_2, \ddot{y}_2) \;&=\; \int 1_{\left\{ \sum_{s=1}^2 \dot{l}_s \left(\dot{y}_s - \frac{1}{2n_s^{1/2}} \ddot{y}_s \dot{y}_s \right) \geq x \right\}} \psi_2(\mathbf{w}_2^n, \dot{y}_2, \ddot{y}_2) 1_{\{l_2 \in \mathcal{A}_{2,c}^n\}} \nu_{c_2}^n(dl_2, d\mathbf{w}_2^n) \\
&\qquad \times q\big((l_1, \sigma_1^{-1} y_1), dc_2\big) \psi_1(\mathbf{w}_1^n, \dot{y}_1, \ddot{y}_1) 1_{\{l_1 \in \mathcal{A}_{1,c}^n\}} \nu_{c_1}^n(dl_1, d\mathbf{w}_1^n).
\end{aligned}$$

To carry out the integration of $\widehat{f}_0^n$, we apply the importance sampling. Let $x > 0$. Prepare constants $a_s \in \mathbb{R}^2$ for $s \in \mathbb{S}$. For example,

$$a_s \;=\; \big(\sum_{s \in \mathbb{S}} |\dot{l}_s|^2 \big)^{-1} \begin{pmatrix} \dot{l}_s^{(1)} \\ \dot{l}_s^{(2)} \end{pmatrix},$$

where $\dot{l}_s = (\dot{l}_s^{(1)}, \dot{l}_s^{(2)})$. Suppose that $(\xi_s^i)_{i \in \mathbb{N}}$ is a three-dimensional i.i.d. sequence such that

$$\xi_s^i \;=\; \begin{pmatrix} \dot{\xi}_s^i \\ \ddot{\xi}_s^i \end{pmatrix} \sim N_3\left(\mu_s,\ \Sigma_s\right), \qquad \mu_s \;=\; \begin{pmatrix} x a_s \\ 0 \end{pmatrix},$$

for $s \in \mathbb{S}$, and that $(\xi_i^1)_{i\in\mathbb{N}}$ and $(\xi_i^2)_{i\in\mathbb{N}}$ are independent. The matrix Σ_s is a positive-definite matrix bigger than the covariance matrix of the normal distribution $N_3(0, \mathsf{V}_s)$ of the principal part of ψ_s. For example, $\Sigma_s = p\,\mathsf{V}_s$ for some constant $p > 1$. We approximate $\widehat{f}_0^n$ by

$$\mathcal{I}_N \quad = \quad \frac{1}{N}\sum_{i=1}^{N} \frac{\mathcal{J}(\dot{\xi}_1^i, \ddot{\xi}_1^i, \dot{\xi}_2^i, \ddot{\xi}_2^i)}{\phi(\dot{\xi}_2^i, \ddot{\xi}_2^i; \mu_2, \Sigma_2)\phi(\dot{\xi}_1^i, \ddot{\xi}_1^i; \mu_1, \Sigma_1)}$$

with large N. In the case of (98), what should be computed is

$$\mathcal{I}_{N,c_2} \quad = \quad \frac{1}{N}\sum_{i=1}^{N} 1_{\left\{\sum_{s=1}^{2} i_s\left(\dot{\xi}_s^i - \frac{1}{2n_s^{1/2}}\ddot{\xi}_s^i\dot{\xi}_s^i\right)\geq x\right\}} 1_{\left\{(-1)^{c_2+1}\ddot{i}_1\dot{\xi}_1^i > 0\right\}} \frac{\psi_2(\dot{\xi}_2^i, \ddot{\xi}_2^i)\psi_1(\dot{\xi}_1^i, \ddot{\xi}_1^i)}{\phi(\dot{\xi}_2^i, \ddot{\xi}_2^i; \mu_2, \Sigma_2)\phi(\dot{\xi}_1^i, \ddot{\xi}_1^i; \mu_1, \Sigma_1)}$$

for each $c_2 \in \mathfrak{C}_2$. This computation is feasible much faster than the raw simulation of the batched bandits model. It should be remarked that our method based on the asymptotic expansion is a kind of invariance principle in that the approximation formulas depend on the first several moments estimable from the data, while simulation methods need complete information about the distribution of the error.

The measures $\widetilde{\Psi}^n_{s,\mathsf{p},\mathbf{w}_s^n}(d\dot{y}_s, d\ddot{y}_s)$ involve the moments of $\dot{\epsilon}_j$ in the correction terms. In practice, we substitute them by their estimated value. It is possible to use a reduced formula that neglects the terms having the moments of $\dot{\epsilon}_j$ of order larger than four because these moments appear in higher order terms than the first order asymptotic expansion.

D Additional Numerical Results

In this section, we provide additional simulations to verify how our methods numerically perform. We consider two stage batched bandit experiments. In addition to canonical applications like exploratory and confirmatory clinical trials, from a theoretical point of view,adding approximately gaussian(but not completely gaussian)test statistic stagewise will induce a stage wise central limit theorem, which is not incorporated in our theory. Hence to check whether our theory is valid, fixing the number of batches to two, and varying the sample size and the noise distribution is the best way. In the first subsection, we will vary the number of sample size from 10, 30 to 50, and see how well normal, and the asymptotic expansion approximate the true underlying distribution[6] In the next subsection, we will fix the sample size, and change the noise distribution far from normal, to exactly normal. If our theory is correct, our asymptotic expansion formula should perform better relatively better, when the sample size is small and the noise distribution is far from normal.

D.1 Varying sample size

In this section, we consider the underlying distribution to be $\Gamma 2, 1$, and change the sample size from 10, 30, 50 for each stage. (In total of two stages, they should have 20,60,100 samples to create the test statistic) As we can see from the table and plots, with sample size around 10 or 30, the underlying distribution is clearly far from normal. On the other hand, the quantile calculated based on our asymptotic expansion formula seems to correct about half of the approximation error, and it is usually uniformly better relative to the normal quantiles. Also , within the asymptotic expansion, the difference between the second order and the first order asymptotic expansion is evident especially in small sample size, but the difference becomes marginal around 50 samples per batch.

D.2 varying noise distribution

Next, we will fix the sample size to 20 per batch, and vary the noise distribution. Varying the Gamma distribution is an easy way to manipulate the skewness, kurtosis of the underlying distribution. We compare the $\Gamma(2, 1)$(the chi -squared distribution)with degree one) which clearly has strong

[6]We calculated the 'true' underlyin distribution by monte carlo with 1000000 times, and standard errors < 0.0001 . We avoided putting the error bars in the plot, which will hinder the (sometimes)subtle difference between the methods. Our codes for the numerical experiments are available here https://1drv.ms/u/s!AiBMsb9GEfTahJMbGoaVnGxza6AsHg?e=8uGmxy

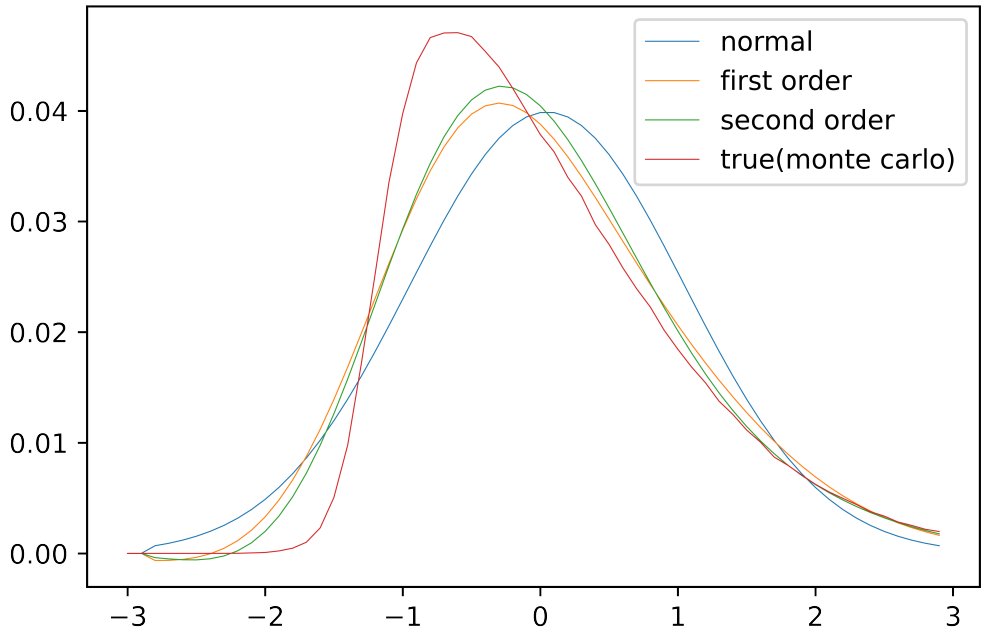

Figure 3: distribution of $\Gamma(1, 2)$, 10 sample per batch

Table 1: Comparison between the quantiles: $\Gamma(1, 2)$, 10 sample per batch

Method\Probability	0.025	0.050	0.950	0.975
Monte Carlo	-1.36	-1.24	1.91	2.41
2nd order Asymptotic expansion	-1.67	-1.48	1.87	2.29
1st order Asymptotic expansion	-1.61	-1.43	1.86	2.35
Normal approximation	-1.96	-1.64	1.64	1.96

Standardized $\Gamma(1, 2)$as noise distribution, 10 samples per batch, epsilon-greedy with clipping constraint 0.1

Table 2: Comparison between the quantiles: 30 sample per batch

Method\Probability	0.025	0.050	0.950	0.975
Monte Carlo	-1.59	-1.4	1.83	2.24
2nd order Asymptotic expansion	-1.81	-1.55	1.73	2.12
1st order Asymptotic expansion	-1.82	-1.56	1.74	2.11
Normal approximation	-1.96	-1.64	1.64	1.96

Standardized $\Gamma(1, 2)$as noise distribution, 30 samples per batch, epsilon-greedy with clipping constraint 0.1

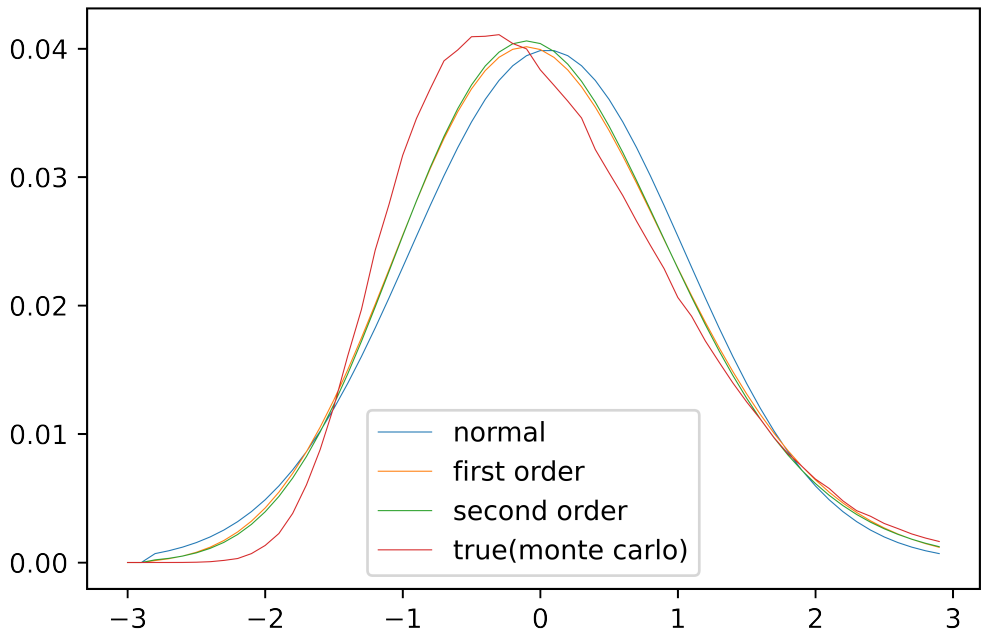

Figure 4: distribution of $\Gamma(1, 2)$, 30 sample per batch

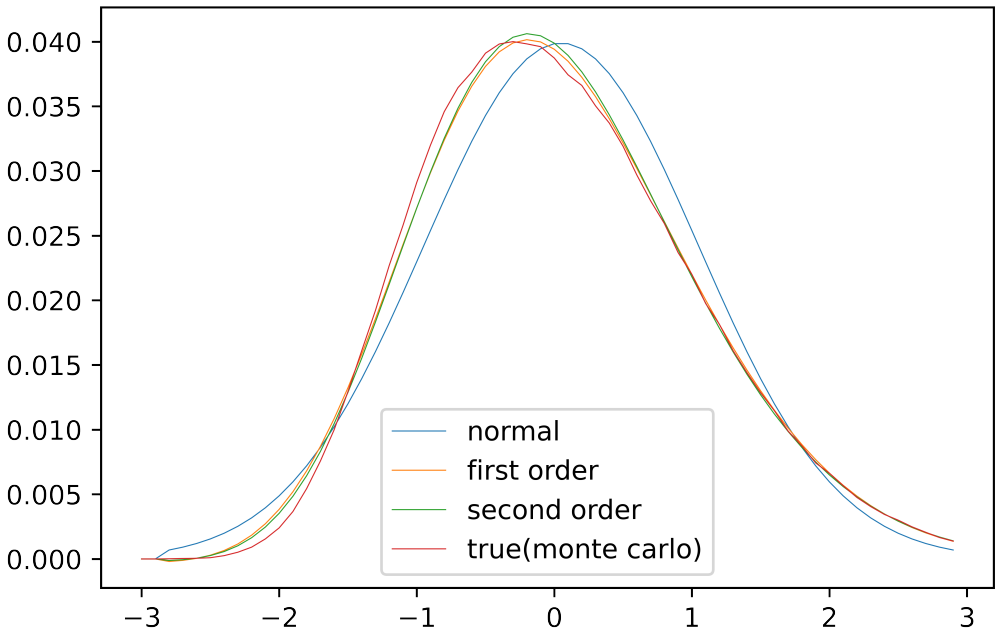

Figure 5: distribution of $\Gamma(1, 2)$, 50 sample per batch

Table 3: Comparison between the quantiles: 50 sample per batch

Method\Probability	0.025	0.050	0.950	0.975
Monte Carlo	-1.67	-1.47	1.79	2.19
2nd order Asymptotic expansion	-1.73	-1.51	1.79	2.18
1st order Asymptotic expansion	-1.75	-1.52	1.78	2.17
Normal approximation	-1.96	-1.64	1.64	1.96

Standardized $\Gamma(1,2)$as noise distribution, 50 samples per batch, epsilon-greedy with clipping constraint 0.1

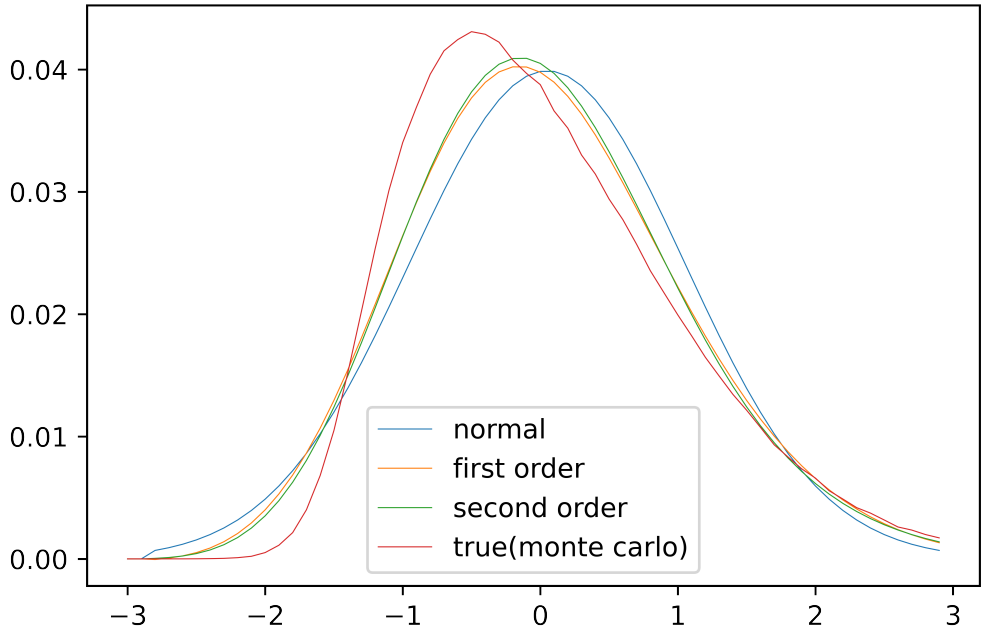

Figure 6: distribution of $\Gamma(1,2)$, 20 sample per batch

skewness and kurtosis, then Ga(2,2,) which is slightly skewed , but closer to normal, and finally the normal distribution. If our theory is correct, in the case of the normal distribution, the higher order terms in our asymptotic expansion density should be zero, so there should be no difference between the normal and higher order approximation.

As in Table 2, we can see that, as the true underlying distribution of the test statistic is farther from normal, in the order of $\Gamma(2,1)$, $\Gamma(2,2)$, and $N(0,1)$, the asymptotic expansion performs better. We can see this clearly in the case of $\Gamma(2,1)$. In the case of $\Gamma(2,2)$, the normal approximation is not that bad, but still asymptotic expansion seems to have uniform improvement over the normal approximation.

Table 4: Comparison between the quantiles: $\Gamma(1,2)$ 20 sample for each batch distributions

Method\Probability	0.025	0.050	0.950	0.975
Monte Carlo	-1.51	-1.36	1.85	2.3
2nd order Asymptotic expansion	-1.76	-1.52	1.76	2.17
1st order Asymptotic expansion	-1.78	-1.54	1.77	2.15
Normal approximation	-1.96	-1.64	1.64	1.96

Standardized $\Gamma(1,2)$as noise distribution, 20 samples per batch, epsilon-greedy with clipping constraint 0.1

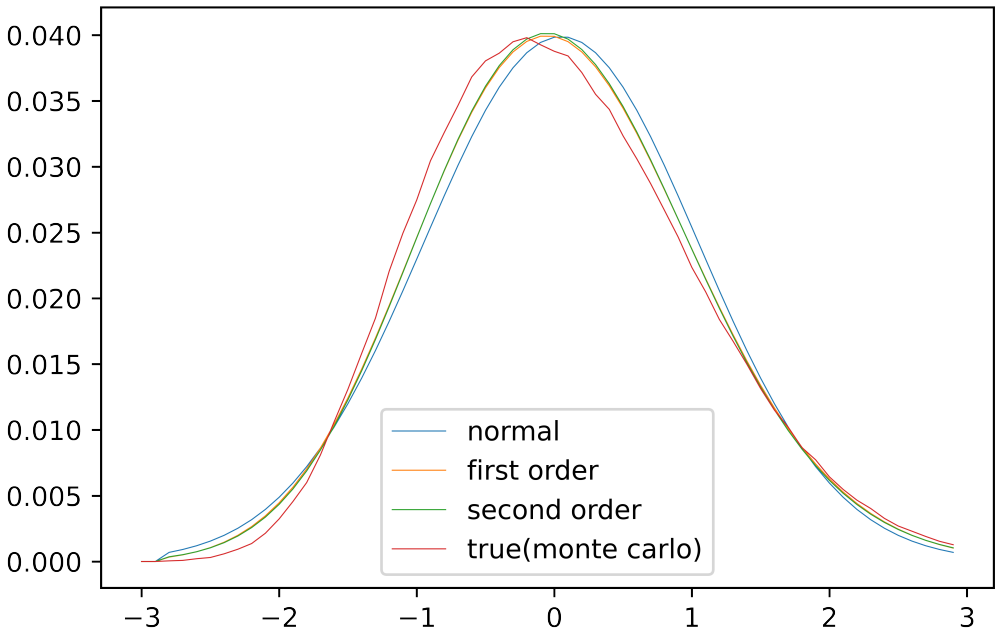

Figure 7: distribution of $\Gamma(2,2)$, 20 sample per batch

Table 5: Comparison between the quantiles: $\Gamma(2,2)$ 20 sample for each batch distributions

Method\Probability	0.025	0.050	0.950	0.975
Monte Carlo	-1.73	-1.51	1.75	2.14
2nd order Asymptotic expansion	-1.86	-1.58	1.71	2.06
1st order Asymptotic expansion	-1.87	-1.58	1.78	2.06
Normal approximation	-1.96	-1.64	1.64	1.96

Comparison between the quantiles: $\Gamma(2,2)$ 20 sample for each batch distributions

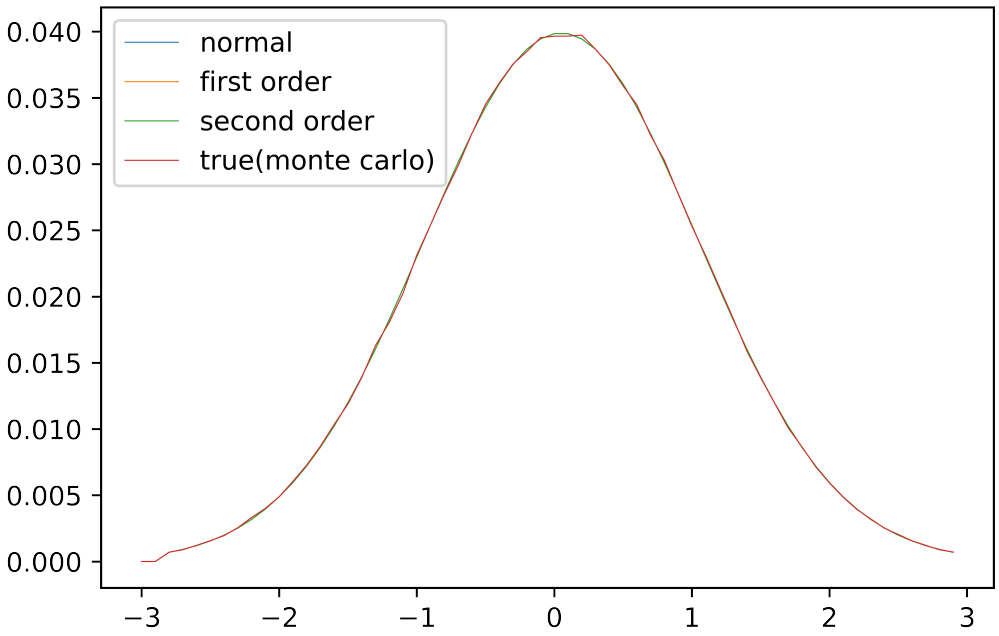

Figure 8: distribution of $N(0,1)$, 20 sample per batch

Table 6: Comparison between the quantiles: $N(0,1)$ 20 sample for each batch distributions

Method\Probability	0.025	0.050	0.950	0.975
Monte Carlo	-1.97	-1.64	1.64	1.96
2nd order Asymptotic expansion	-1.96	-1.64	1.64	1.96
1st order Asymptotic expansion	-1.96	-1.65	1.64	1.96
Normal approximation	-1.96	-1.64	1.64	1.96

Standardized $N(0,1)$as noise distribution, 20 samples per batch, epsilon-greedy with clipping constraint 0.1