Published in final edited form as:
J Stat Plan Inference. 2010 November ; 140(11): 3505–3519. doi:10.1016/j.jspi.2010.05.022.

Compound adaptive GPU design for clinical trials

Ao Yuan*, **Paul Bezandry**, and **George Bonney**
Howard University, Washington, DC 20059, USA

Paul Bezandry: pbezandry@Howard.edu; George Bonney: ge_bonney@howard.edu

Abstract

In clinical trials, several competing treatments are often carried out in the same trial period. The goal is to assess the performances of these different treatments according to some optimality criterion and minimize risks to the patients in the entire process of the study. For this, each coming patient is allocated sequentially to one of the treatments according to a mechanism defined by the optimality criterion. In practice, sometimes different optimality criteria, or the same criterion with different regimes, need to be considered to assess the treatments in the same study, so that each mechanism is also evaluated through the trail study. In this case, the question is how to allocate the treatments to the incoming patients so that the criteria/mechanisms of interest are assessed during the trail process, and the overall performance of the trial is optimized under the combined criteria or regimes. In this paper, we consider this problem by investigating a compound adaptive generalized Pólya urn design. Basic asymptotic properties of this design are also studied.

Keywords

Adaptive design; Clinical trials; Compound generalized Pólya urn; Optimal design

1. Introduction

1.1. Brief review of clinical trails and the Pólya urn design

In clinical trials, there are often several competing treatments to be assessed during the trial process while patients come sequentially over the long period of study. Based on the then current clinical knowledge of the treatments, each coming patient is allocated to one of them according to some defined mechanisms, so that the overall treatment loss is minimized by the given optimality criteria. However, there are cases in which the treatments needs to be assessed under different criteria and the criteria themselves are to be assessed during the trial process, thus the patients' allocation needs to be determined by the different random mechanisms, in which multiple criteria with different regimes are considered in the same study. In practice, often the same treatments under different mechanisms results differently, thus may require different criterion for each mechanism. As each treatment has its strength and weakness, thus we want to assess the treatments under each of the criterion and allocate the patients so as to minimize the overall risk. The results under each criterion/mechanism will be analyzed separately so that each of them is to be evaluated. As the treatments performances unknown in prior, and are gradually learned through the trial process, to minimize the overall loss of the trial, a compromised criterion between the two is desirable. As the compound criterion depends on the unknown performances of all the treatments, it cannot be determined in prior, instead, it is implemented along the trial process. In this case,

*Corresponding author. ayuan@howard.edu (A. Yuan).

the question is how to allocate the coming patients so that the allocation is optimized under the compound criterion. Here, we attempt to explore this problem by using a compound generalized Pólya urn design, and will come back at the problem in the Application section. For this, we first review briefly some commonly used methods in clinical trials. Some of these methods apply to either discrete response, continuous response, or both. The generalized Pólya urn (GPU) is a well-known design in clinical trial; there are many extensive results for it. Many of the existing models are applied to the discrete responses, and are not optimally adaptive.

The adaptive design uses accumulating data to update aspects of the study as it continues without undermining the validity and integrity of the trial (Hayre, 1979; Melfi and Page, 1998; Jennison and Turnbull, 2000; Hu and Rosenberger, 2003; Hu and Zhang, 2004, among others). A comprehensive review of works in this filed can be found in Hu and Rosenberger (2006). The purpose of optimal design is to achieve some targeting objective criteria for the allocation proportions (Eisele, 1994; Eisele and Woodroofe, 1995; Rosenberger et al., 2001; Hu et al., 2006; Hu et al., 2009; for example). Recently, Zhang et al. (2006), thereafter ZHC, proposed a sequential estimation-adjusted urn model, Yuan and Chai (2008) independently of ZHC, thereafter YC, studied an adaptive GPU design, which has some similarity to that of ZHC. This design was studied further in Yuan (2008) for the cases of delayed response, staggered/censored entry, heterogeneity and longitudinal/repeated observations. These methods optimize any given target functional of the trial distribution and are applicable to both discrete and continuous responses.

1.2. An exposition of a two-treatment clinical trial

To help understanding of our motivation, consider a simple example. Suppose we have two treatments to investigate in the clinical trial. The success rate of the two treatments are p_1 and p_2 respectively, $N_1(n)$ and $N_2(n)$ are the numbers of patients allocated the each treatment at time n. The commonly used Neyman allocation (Melfi and Page, 1998) is designed to maximize the power of detecting the difference $p_1 - p_2$ of mean performances of the trials, which leads to the allocation proportion $N_1(n)/N_2(n) \rightarrow \sqrt{p_1(1-p_1)}/\sqrt{p_2(1-p_2)}$. While the criterion in Rosenberger et al. (2001) is to minimize the expected treatment failure and leads to the proportion $N_1(n)/N_2(n) \rightarrow \sqrt{p_1/p_2}$. The two criteria give very different allocation results. When the treatment differences are small, the former will be better in terms of distinguishability, thus smaller sample size will be needed in the study, which is much desired since patients are cost to get in practice. But also it may lead to more life losses which will be of grave consequence as the subjects in study are human beings. The second criterion can lead to less treatment losses when the differences are relatively significant but may not have the desirable power to detect the differences between the treatments. We will see in the application section that, as a result of optimizing the two criteria simultaneously with our method, assume $p_1 > p_2$, the compound method will give the allocation proportion $((1-p_1+p_2)/(\sqrt{p_1(1-p_1)}+\sqrt{p_2(1-p_2)}))(\sqrt{p_1(1-p_1)},\ \sqrt{p_2(1-p_2)})+((p_1-p_2)/(\sqrt{p_1}+\sqrt{p_2}))(\sqrt{p_1},\ \sqrt{p_2})$. This proportion is a combination/compromise of the two, and is more robust than either one of two criterion used along, in that it keeps much of the power in detecting the difference while maintain less treatment losses.

In this paper, to assess the treatments by different criteria in the same trial and keep the advantages of each trail and avoid their weakness, we propose and study a compound version of the design in YC. The rest part is organized as follows. In Section 2, we describe the existing adaptive sequential GPU design of YC. Next, we present compound versions

with and without adaptive features and their basic asymptotic properties in Section 3, and an illustrative application in Section 4. The relevant technical proofs are given in the Appendix.

2. Description of the optimal adaptive GPU design

We first describe the adaptive optimal GPU design (YC). Suppose there are k treatments under study, and a Pólya urn with some initial components, corresponding to records of patient assignments to these treatments $\mathbf{X}_0 = (X_{0,1},\ldots,X_{0,k})$. Here, the components of the urn can be discrete, real valued or of mixed type. Let $\mathbf{r}_n$ be the (multiple) response of the n-th patient under one of the treatments and $f(\mathbf{r}_n)$ be the summary score for this response. Without loss of generality we assume $0 \le f(\cdot) < \infty$. Let A_i be the event that the study is under treatment i, $\mu_i = E(f(\mathbf{r}_1)|A_i)$ be the expected performance, or success rate, of the i-th treatment, and $\sigma_i^2 := Var(f(\mathbf{r}_1)|A_i) < \infty$ be its variance ($i = 1,\ldots,k$). Set $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k)$ and $\sigma = (\sigma_1^2, \ldots, \sigma_k^2)$. For any vector $\mathbf{x} = (x_1,\ldots,x_k)$, denote $|\mathbf{x}| = \sum_{i=1}^{k} x_i$ and $\{\mathbf{x}\} = \mathrm{diag}(\mathbf{x})$, the diagonal matrix for $\mathbf{x}$. A vector $\mathbf{x}$ is *normalized* if $|\mathbf{x}| = 1$. At time n, the urn composition is $\mathbf{X}_n = (X_{n,1}, \ldots, X_{n,k})$, the total number of patients assigned to treatment i at time n is $N_i(n)$, and denote $\mathbf{N}(n) = (N_1(n),\ldots,N_k(n))$. In the GPU model, to assign the entering n-th patient to one of the treatments, a random variable is drawn from the multinomial distribution with probabilities $\mathbf{X}_n/|\mathbf{X}_n|$. If it is type i, the patient is assigned to the i-th treatment, a random vector of masses $\boldsymbol{\xi}_i$ is added to the urn compositions, and the response $\mathbf{r}_n$ is used to update the estimate of $\boldsymbol{\mu}$ in the next step. Let $\boldsymbol{\xi}_i = (\xi_{i1}, \ldots, \xi_{ik})$ be the increments to the urn composition given the patient is assigned treatment i, $\boldsymbol{\xi} = (\xi_{ij})_{i,j=1,\ldots,k}$ be the matrix representation of the ξ_{ij}'s, and for each n, $\boldsymbol{\xi}_n$ is an i.i.d. version of $\boldsymbol{\xi}$. To simplify the expressions of the asymptotic variances to be derived later, we assume throughout this article that $\boldsymbol{\xi}$ is independent of the response observations. The random vector $\boldsymbol{\xi}_i$ is termed the adding rule, and $\boldsymbol{v} = E(\boldsymbol{\xi}) = (v_{ij})$ the design matrix with $v_{ij} = E(\xi_{ij})$ (known). The first eigenvalue λ of the design matrix and its normalized first left (row) eigenvector $\mathbf{v}$ play a key role in the asymptotic properties of the GPU design. Many authors (for instance Athreya and Karlin, 1967; Gouet, 1997; Janson, 2004) studied asymptotic properties of $\mathbf{X}_n$ and $\mathbf{N}(n)$ and proved that

$$(\mathbf{X}_n/|\mathbf{X}_n|, \mathbf{N}(n)/n) \to (\mathbf{v}, \mathbf{v}) \ \ (a.s.)$$

For a comprehensive review in this field, see Rosenberger and Lachin (1993) and other related recent papers.

In YC, the optimal adaptive GPU is studied. The idea is to choose a vector $\mathbf{v} = (v_1,\ldots,v_k)$ for a given λ and a specified object functional $G(\cdot, \cdot)$ according to the optimality criterion

$$\mathbf{v} = \underset{\mathbf{v}\in\nabla}{\arg\sup}\, G(\mu, \mathbf{v}(\mu)),\ \nabla = \{\mathbf{v}(\cdot) = (v_1(\cdot), \ldots, v_k(\cdot)), v_j(\cdot) > 0 (1 \le j \le k), |\mathbf{v}| = 1.\},$$

and then construct special adding rule $\boldsymbol{\zeta}'_{ni}$ below in such a way that the design matrix has the above vector $\mathbf{v}(\cdot)$ as its first left eigenvector and λ as the first eigenvalue. It turns out that the urn composition will be asymptotically optimal as quantified by the given by $\mathbf{v}_1(\cdot)$". More precisely, let I_{ji} be the indicator that the j-th patient is assigned to treatment i. Since $\boldsymbol{\mu}$ is unknown, we plug in its current estimate $\boldsymbol{\mu}_n$, with

$$\mu_{ni}=\frac{\sum_{j=1}^{n} f(\mathbf{r}_j)I_{ji}}{\sum_{j=1}^{n} I_{ji}},\ \ (1 \le i \le k)$$

and in place of ξ_{ni}, we add $\boldsymbol{\zeta}_{ni}$ mass to the urn if the coming patient is assigned treatment i, where

$$\zeta_{n,ij}=\frac{m_{ij}(\mu_{n-1})\xi_{n,ij}}{v_{ij}},\ \ (1 \le i,\ j \le k).$$

The matrix $\mathbf{M} = (m_{ij}(\boldsymbol{\mu}_{n-1}))$ will be specified later. Here the adding rule $\zeta_{n,ij}$'s are functions of the responses via $\boldsymbol{\mu}_{n-1}$, and we keep $\xi_{n,ij}/v_{ij}$ in the rule to make it general and flexible, as the literature for this topic indicated. The matrix $\mathbf{M} = (m_{ij})$ is constructed according to the above optimality criterion, so that the ultimate urn composition will achieve this criterion, which is the goal of the study. Let $\boldsymbol{\zeta}_n = \boldsymbol{\zeta}_n(\boldsymbol{\mu}_{n-1}) = (\zeta_{n,ij})_{i,j = 1,\ldots,k}$, $\zeta_{ij} = m_{ij}(\boldsymbol{\mu})\xi_{ij}/v_{ij}$, and $\bar{\mathbf{M}} = \bar{\mathbf{M}}(\boldsymbol{\mu}) = \mathbf{M}(\boldsymbol{\mu}) - \mathbf{1}'\mathbf{v}(\boldsymbol{\mu})$, $\boldsymbol{\zeta} = \boldsymbol{\zeta}\,(\boldsymbol{\mu}, \xi_1)$, $\boldsymbol{\Sigma}_1 = \{\mathbf{v}\} - \mathbf{v}'\mathbf{v}$, $\boldsymbol{\Sigma}_2 = E[(\boldsymbol{\zeta} - \mathbf{M})'\{\mathbf{v}\}(\boldsymbol{\zeta} - \mathbf{M})]$ and $\boldsymbol{\Sigma}_3 = \{\mathbf{v}\sigma^2\}$. Define $\mathbf{t}_j = \mathbf{v}\partial\mathbf{M}(\mathbf{x})/\partial x_j|_{\mathbf{x} = \boldsymbol{\mu}}$ and $\mathbf{T}=(\mathbf{t}_1'/v_1,\ldots,\mathbf{t}_k'/v_k)'$.

The main goal for an optimal design is to achieve, asymptotically, some specified optimality criterion for the allocation proportions $\mathbf{N}(n)/n$. Let $\mathbf{1}$ be a row vector of 1's with proper length. The asymptotic properties of the urn in ZHC and YC are obtained under the following conditions:

(B0) $0 < v_{ij} < \infty$, $(1 \le i,j \le k)$

(B1) $\mathbf{M1}' = \gamma\mathbf{1}'$ for some $\gamma > 0$

(B2) $\zeta_{ij}(\cdot,\cdot)$ is continuous at $(\boldsymbol{\mu},\mathbf{y})$ for any $\mathbf{y}$ and there is an $r > 2$ such that

$$E\|\boldsymbol{\xi}_1\|^r<\infty \ \text{ and } \sup_n E(\|\boldsymbol{\zeta}_n\|^r|\mathcal{F}_{n-1})<\infty,\ \text{ a.s.}$$

(B3) $\mathbf{M}(\boldsymbol{\mu}) > \mathbf{0}$ is bounded, differentiable, and there is a $\delta > 0$ such that

$$\mathbf{M}(\mathbf{y}) - \mathbf{M}(\mu)=\sum_{j=1}^{k}\frac{\partial\mathbf{M}(\mathbf{y})}{\partial y_j}\Big|_{\mathbf{y}=\mu}(y_j-\mu_j)+\mathrm{O}(\|\mathbf{y}-\mu\|^{1+\delta}),\ \text{ as } \mathbf{y}\to\mu.$$

The following result is proven in (YC).

Under (B0)–(B2) (for some r > 1), we have

i. $\boldsymbol{\mu}_n \to \boldsymbol{\mu}$, *a.s.*, $\sup_{\mathbf{x}\in S(\mathbf{F})}\|\mathbf{F}_n(\mathbf{x}) - \mathbf{F}(\mathbf{x})\| \to 0$, *a.s.*

ii. $(\mathbf{X}_n(\boldsymbol{\mu}_{n-1})/n, \mathbf{N}(n)/n) \to (\lambda\mathbf{v}(\boldsymbol{\mu}), \mathbf{v}(\boldsymbol{\mu}))$ *a.s.*

iii. $\sum_{i=1}^{k}\sum_{j=1}^{n} f(\mathbf{r}_j)I_{ji}/n \to \mathbf{v}'(\mu)\mu,\ a.s.$

Under (B0)–(B2), for some given matrices $\boldsymbol{\Omega}_\mu$ and $\boldsymbol{\Omega}_F$, we have

$$\sqrt{n}(\mu_n - \mu) \xrightarrow{d} N(\mathbf{0}, \mathbf{\Omega}_\mu), \quad \sqrt{n}(\mathbf{F}_n(\mathbf{x}) - \mathbf{F}(\mathbf{x})) \xrightarrow{d} N(\mathbf{0}, \mathbf{\Omega}_F), \quad \mathbf{x} \in S(\mathbf{F}).$$

Under (B0)–(B3), If $\lambda_1 > 2\mathrm{Re}(\lambda_2)$, then for some given $\mathbf{\Omega}$,

$$\sqrt{n}\left(\frac{\mathbf{X}_n(\mu_{n-1})}{n} - \lambda\mathbf{v}(\mu), \frac{\mathbf{N}(n)}{n} - \mathbf{v}(\mu)\right)' \xrightarrow{d} N(\mathbf{0}, \mathbf{\Omega}).$$

These results depict the asymptotic optimality achieved by the allocation proportions $\mathbf{N}(n)/n$, as quantified by the given $\mathbf{v}_1(\boldsymbol{\mu})$, and give its asymptotic normality, in which the asymptotic covariance matrix $\mathbf{\Omega}$ is an important quantity to evaluate the performance of the design.

3. Compound GPU design

In this section, we study the compound version of the optimal GPU model discussed in Section 2 in which r different allocation mechanisms are considered in clinical trial study. Before the outline of the compound GPU design, let us first introduce the notations we use in the sequel. Let B_s be the event that the underlying patient is assigned to mechanism s and A_j be the event that the patient is allocated to treatment j. Denote by $\mathbf{r}_n|B_s$ the (multiple) response of the n-th patient under one of the treatments in mechanism s and by $f(\mathbf{r}_n)|B_s$ the summary score for this response. Let $F_i^{(s)}(\cdot)$ be the distribution function of $f(\mathbf{r}_n)|B_s,A_i$, which describes the conditional performance of the scores $f(\mathbf{r}_n)|B_s$'s on the i-th treatment, $\mu_i^{(s)}=E[f(\mathbf{r}_1)|A_i, B_s]$ be the expected performance, or success rate, of the i-th treatment in mechanism s, and $\mu^{(s)}=(\mu_1^{(s)}, \ldots, \mu_k^{(s)})$, $\mathbf{F}^{(s)}=(F_1^{(s)}, \ldots, F_k^{(s)})$. In practice, we often have $\boldsymbol{\mu}^{(s)} = \boldsymbol{\mu}$ and $\mathbf{F}^{(s)} = \mathbf{F}$ ($s = 1,\ldots,r$), where $\mathbf{F} = (F_1,\ldots,F_k)$ with F_j is the distribution of $f(\mathbf{r}_n)|A_j$ and $\boldsymbol{\mu}$ is the mean of $\mathbf{F}$. Also, $f(\mathbf{r}_n)|\ B_s$ denotes that individual n is assigned to the mechanism s (here we use the notation $|\ B_s$ to emphasize the condition that the underlying patient is assigned to mechanism s). Without loss of generality we assume that each mechanism is characterized by an optimality criterion $G_s(\cdot)$. In turn, such an optimality can be implemented in a design matrix and summarized by the normalized and non-negative first left row eigenvector $\mathbf{v}_s$ ($s = 1, \ldots, r$) of the design matrix as

$$\mathbf{v}_s=\underset{\mathbf{v}\in\nabla}{\operatorname{argsup}}\, G_s(\mathbf{F}), \quad \nabla=\{\mathbf{v}(\cdot)=(v_1(\cdot), \ldots, v_k(\cdot)), v_j(\cdot)>0(1 \le j \le k), |\mathbf{v}|=1\}.$$

Each of these mechanisms will be randomly assigned in the clinical trial study according to another criterion $H(\cdot)$ and characterized by another normalized first left eigenvector $\mathbf{u} = (u_1, \ldots,u_r)$ as

$$\mathbf{u}=\underset{\mathbf{u}\in\Delta}{\operatorname{argsup}}\, H(G_1(\mathbf{F}), \ldots, G_r(\mathbf{F}), \quad \Delta=\{\mathbf{u}(\cdot)=(u_1(\cdot), \ldots, u_r(\cdot)), u_j(\cdot)>0(1 \le j \le r), |\mathbf{u}|=1.\}.$$

Here, we assume that the respective design matrices denoted by $\mathbf{M}^{(s)}$ and $\mathbf{Q}$ depend on $\mathbf{F}$ through $\boldsymbol{\mu}$. The requirement $v_j(\cdot) > 0$ $(1 \le j \le k)$ (resp. $u_j(\cdot) > 0$ $(1 \le j \le r)$) is to guarantee that $\mathbf{v}_s$ (resp. $\mathbf{u}$) is indeed the first right eigenvector, and that the design matrix $\mathbf{M}^{(s)}=\mathbf{M}^{(s)}(\mathbf{F})=(m_{ij}^{(s)}(\mu))_{k\times k}$, (resp. $\mathbf{Q} = \mathbf{Q}(\mathbf{F}) = (q_{ij}(\boldsymbol{\mu}))_{r\times r}$) can be constructed in the same

manner as in (YC) using $\mathbf{v}_s$ and its corresponding eigenvalue λ_s (resp. $\mathbf{u}$ and its corresponding eigenvalue γ).

Associated with mechanism s, there is urn s with composition $\mathbf{X}_n^{(s)}=(X_{n,1(s)},\ldots,X_{n,k(s)})$ at time n ($s = 1,\ldots,r$), and with the criterion $\mathbf{Q}$ there is an urn with composition $\mathbf{Y}_n = (Y_{n,1}, \ldots, Y_{n,r})$ at time n. Each of these urns has an adding rule. $\boldsymbol{\xi}^{(s)}=(\xi_{ij}^{(s)})_{k\times k}$ is referred as the adding rule corresponding to $\mathbf{M}^{(s)}$ with mean $E\boldsymbol{\xi}^{(s)}=(v_{ij}^{(s)})_{k\times k}$ and $\boldsymbol{\eta} = (\eta_{ij})_{r\times;r}$ as the adding rule corresponding to $\mathbf{Q}$ with $E\boldsymbol{\eta} = (\delta_{ij})_{r\times r}$. Also, each of the urns is assumed to have some non-negative initial compositions at $n = 0$. The overall urn composition is $\mathbf{X}_n=\sum_{s=1}^{r}\mathbf{X}_n^{(s)}$, which is the composition of the k treatments at time n. Let $I_{m(s)}$ be the indicator that m-th patient is in mechanism s and $N_i^{(s)}(n)=\sum_{m=1}^{n}I_m^{(s)}I_{mi}$ be the total number of patients assigned to treatment i in mechanism s at time n. Let $\mathbf{N}^{(s)}(n)=(N_1^{(s)}(n),\ldots,N_{k(s)}(n))$ be the numbers of patients allocated to each treatment in mechanism s at time n, $N_i(n)=\sum_{s=1}^{r}N_i^{(s)}(n)\ (i=1,\ldots,k)$, $n^{(s)} = |\mathbf{N}^{(s)}(n)|$, and $\mathbf{n}(n)=(n^{(1)},\ldots,n_{(r)})$. The total allocated numbers is $\mathbf{N}(n)=\sum_{s=1}^{r}\mathbf{N}^{(s)}(n)=(N_1(n),\ldots,N_k(n))$, and $n = |\mathbf{N}(n)|$. The empirical estimates of $\boldsymbol{\mu}$ and $\mathbf{F}$ are given by $\boldsymbol{\mu}_n = (\mu_{n,1}, \ldots, \mu_{n,k})$ and $\mathbf{F}_n = (F_{n,1},\ldots,F_{n,k})$, with

$$\mu_{ni}=\sum_{s=1}^{r}(n^{(s)}/n)\mu_{ni}^{(s)} \text{ and } F_{ni}(x_i)=\sum_{s=1}^{r}(n^{(s)}/n)F_{ni}^{(s)}(x_i),\ \ (i=1,\ldots,k)$$

where $\mu_{ni}^{(s)}=\sum_{j=1}^{n}I_j^{(s)}I_{ji}f(\mathbf{r}_j)/\sum_{j=1}^{n}I_j^{(s)}I_{ji}$, and $F_{ni}^{(s)}(x_i)=\sum_{j=1}^{n}I_j^{(s)}I_{ji}I(f(\mathbf{r}_j)\le x_i)/\sum_{j=1}^{n}I_j^{(s)}I_{ji}$, ($s = 1, \ldots, r$; $i = 1, \ldots, k$). We also denote $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_r)$ with λ_s is the first eigenvalue of the s-th urn, $\mathbf{v}_s = (v_{s1},\ldots,v_{sk})$, and $\mathbf{V} = (v_{ij})_{r\times k}$. Define $\mathbf{u}\odot\boldsymbol{\lambda}\odot\mathbf{V}=\sum_{s=1}^{r}u_s\lambda_s\mathbf{v}_s$, $\mathbf{u}\odot\mathbf{V}=\sum_{s=1}^{r}u_s\mathbf{v}_s$, and for any matrix $\boldsymbol{\Omega}=(\omega_{ij})_{r\times r}\mathbf{V}'\odot\boldsymbol{\Omega}\odot\mathbf{V}=\sum_{1\le i,j\le r}\omega_{ij}\mathbf{v}_i'\mathbf{v}_j$. For $l = k$, r, denote $\mathbf{1}_{l\times l}$ as the $l \times l$ matrix whose elements are all equal to 1. For any matrix $\mathbf{A}$ and vectors $\mathbf{b}$, $\mathbf{c}$, define $\mathbf{a}_i$ the i-th row of $\mathbf{A}$, $\mathbf{a}_i$ the i-th column of $\mathbf{A}$, $\mathbf{bc} = (b_1c_1,\ldots,b_kc_k)$, and $\mathbf{b}/\mathbf{c} = (b_1/c_1,\ldots,b_k/c_k)$.

We now study the compound GPU design and its adaptive version.

3.1. Compound GPU design

In clinical trial study, there are often several competing treatments to be assessed during the trial process and patients entering the study may be very diverse. For example, patients may have different gender, ages, or similar diseases but at different stages. In this case, each treatment may have different effect for these different groups of patients. Here, we consider r regimes and in each regime, there are k treatments. Patients arrive sequentially to be assigned regimes before they are allocated to the treatments. When the n-th patient comes in, we first decide which one of the r regimes should be assigned. For this, we draw a sample from the multinomial distribution ($r,\mathbf{Y}_{n-1}/|\mathbf{Y}_{n-1}|$). If the outcome is s, the patient is assigned to regime s represented by urn s and add the following composition to urn $\mathbf{Y}_{n-1}$

$$\theta_{sj}=\frac{q_{sj}(\mu)\eta_{sj}}{\delta_{sj}}, \quad (j=1,\ldots,r); \ \theta=(\theta_{sj})_{r\times r} \tag{1}$$

as in YC. The urn composition at time n is $\mathbf{Y}_n$. Then we decide which one of the k treatments in mechanism s will be assigned to the n-th patient. For this, we draw a sample from the multinomial distribution $(k, \mathbf{X}_{n-1}^{(s)}/|\mathbf{X}_{n-1}^{(s)}|)$. If the outcome is i, the patient is allocated to treatment i and add the following composition to urn $\mathbf{X}_{n-1}^{(s)}$

$$\zeta_{ij}^{(s)}=\frac{m_{ij}^{(s)}(\mu)\xi_{ij}^{(s)}}{v_{ij}^{(s)}}, \quad (j=1,\ldots,k), \ \boldsymbol{\zeta}^{(s)}=(\zeta_{ij}^{(s)})_{k\times k}. \tag{2}$$

The corresponding urn composition at time n is $\mathbf{X}_n^{(s)}$. In clinical trial study, the asymptotic allocation proportions and the overall urn compositions are of major interest.

As in Section 2, we adopt the following assumptions:

(C0) $0<v_{ij}^{(s)}<\infty$, $(1 \le i,j \le k;\ s = 1, \ldots, r)$ and $0 < \delta_{ij} < \infty$, $(1 \le i,j \le r)$.

(C1) $\mathbf{M}^{(s)}\mathbf{1}_k=\gamma_s\mathbf{1}_k'$ for some $\gamma_s > 0$ $(s = 1,\ldots, r)$; $\mathbf{Q}\mathbf{1}_r'=\gamma\mathbf{1}_r'$ for some $\gamma > 0$. Here, $\mathbf{1}_k = (1,\ldots,1)_k$ and $\mathbf{1}_r = (1,\ldots,1)_r$.

(C3) There is an $r > 2$ such that $E\|\zeta_s\|^r < \infty$ $(s = 1,\ldots,r)$.

Under the design given in (1) and (2), the asymptotic urn compositions and the allocation proportions are characterized by the following theorem.

Theorem 1—Assume (C0)–(C1). Then for the design given in (1) and (2) we have:

i. $$\left(\frac{\mathbf{Y}_n}{n}, \frac{\mathbf{n}(n)}{n}, \frac{\mathbf{N}^{(s)}(n)}{n^{(s)}}\right) \to (\gamma\mathbf{u}, \mathbf{u}, \mathbf{v}_s) \ a.s. \,(s=1,\ldots,r).$$

ii. $$\left(\frac{\mathbf{X}_n}{n}, \frac{\mathbf{N}(n)}{n}\right) \to (\mathbf{u}\odot\boldsymbol{\lambda}\odot\mathbf{V}, \mathbf{u}\odot\mathbf{V}) \ a.s.$$

Now we study the asymptotic distribution of the compound GPU design. Let $\bar{\mathbf{M}}^{(s)} = \mathbf{M}^{(s)} - \mathbf{1}'\mathbf{v}_s$,

$$\textstyle\sum_1^{(s)}=\{\mathbf{v}_s\}-\mathbf{v}_s'\mathbf{v}_s, \ \sum_2^{(s)}=E[(\boldsymbol{\zeta}_s-\mathbf{M}^{(s)})'\{\mathbf{v}_s\}(\boldsymbol{\zeta}_s-\mathbf{M}^{(s)})], \ \sum_3^{(s)}=\{\mathbf{v}_s\sigma^2\}, \ \sum_{23}^{(s)}=E[(\boldsymbol{\zeta}^{(s)}-\mathbf{M}^{(s)})'\{\mathbf{v}_s\}(f(\mathbf{r}_1^{(s)})-\mu^{(s)})], \ \sum^{(s)}=\mathbf{v}_s'\mathbf{e}_s-\mathbf{v}_s'\mathbf{u}$$

, and $\sum_4^{(s)}=(\sigma^2\mathbf{v}_s)'\mathbf{e}_s$, where $\mathbf{e}_s = (0,\ldots,0,1,0,\ldots,0)$ is the s-th unit r-dimensional row vector $(s = 1,\ldots,r)$. Let $\boldsymbol{\Sigma}_1 = \{\mathbf{u}\} - \mathbf{u}'\mathbf{u}$, $\boldsymbol{\Sigma}_2 = E[(\boldsymbol{\theta} - \mathbf{Q})'\{\mathbf{u}\}(\boldsymbol{\theta} - \mathbf{Q})]$, $\boldsymbol{\Sigma}_3 = \{\mathbf{u}\sigma^2\}$, $\boldsymbol{\Sigma}_{23} = E[(\boldsymbol{\theta} - \mathbf{Q})'\{\mathbf{u}\}(f(\mathbf{r}_1) - \boldsymbol{\mu})]$, $\bar{\mathbf{Q}} = \mathbf{Q} - \mathbf{1}'\mathbf{u}$, $\bar{\mathbf{q}}_j$ be the j-th column of $\bar{\mathbf{Q}}$, $\mathbf{h}_s = \mathbf{u}\,\partial\mathbf{Q}(\mathbf{y})/\partial y_s|\mathbf{y} = \boldsymbol{\mu}$ and $\mathbf{H}=(\mathbf{h}_1'/u_1,\ldots,\mathbf{h}_r'/u_r)'$.

For $s = 1,\ldots,r$, define

$$\mathbf{\Lambda}^{\dagger}_{s,1}=\int_1^0\left(\tfrac{1}{x}\right)^{\left(\overline{\mathbf{M}^{(s)}}\right)'}\Sigma_1^{(s)}\left(\tfrac{1}{x}\right)^{\overline{\mathbf{M}^{(s)}}}dx,\ \ \mathbf{\Lambda}^{\diamond}_{s,2}=\int_x^0\left(\tfrac{1}{x}\right)^{\left(\overline{\mathbf{M}^{(s)}}\right)'}\Sigma_2^{(s)}\int_x^1\tfrac{1}{y}\left(\tfrac{y}{x}\right)^{\overline{\mathbf{M}^{(s)}}}dy\,dx,$$
$$\mathbf{\Lambda}^{\sharp}_{s,2}=\int_0^1\left[\int_x^1\tfrac{1}{y}\left(\tfrac{y}{x}\right)^{\overline{\mathbf{M}^{(s)}}}dy\right]'\Sigma_2^{(s)}\left[\int_x^1\tfrac{1}{y}\left(\tfrac{y}{x}\right)^{\overline{\mathbf{M}^{(s)}}}dy\right]dx,\mathbf{\Lambda}^{\dagger}_{s,2}=\int_0^1\left(\tfrac{1}{x}\right)^{\left(\overline{\mathbf{M}^{(s)}}\right)'}\Sigma_2^{(s)}\left(\tfrac{1}{x}\right)^{\overline{\mathbf{M}^{(s)}}}dx.$$

Also, for $i = 1,\ldots,r$ and $j = 1,\ldots,r$, define

$$\mathbf{\Lambda}_{i,j}=\int_0^1\left(\frac{1}{x}\right)^{\left(\overline{\mathbf{M}^{(i)}}\right)'}\mathbf{v}_i'\mathbf{v}_j\left(\frac{1}{x}\right)^{\overline{\mathbf{M}^{(j)}}}dx,\ \ \rho_{ij}=\int_0^1\left(\frac{1}{x}\right)^{\left(\overline{\mathbf{M}^{(i)}}\right)'}\sum\nolimits^{(i)}\left(\frac{1}{x}\right)^{\overline{\mathbf{q}}_j}dx.$$

Finally, define

$$\mathbf{\Lambda}_1^{\dagger}=\int_0^1\left(\frac{1}{x}\right)^{\overline{\mathbf{Q}}'}\sum\nolimits_1\left(\frac{1}{x}\right)^{\overline{\mathbf{Q}}}dx,\ \ \Lambda_2^{\sharp}=\int_0^1\left[\int_x^1\frac{1}{y}\left(\frac{y}{x}\right)^{\overline{\mathbf{Q}}}dy\right]'\sum\nolimits_2\left[\int_x^1\frac{1}{y}\left(\frac{y}{x}\right)^{\overline{\mathbf{Q}}}dy\right]dx.$$

Let $\lambda_{s,2}$ be the second eigenvalue of $\mathbf{M}^{(s)}$ ($s = 1,\ldots,r$), and γ_2 be that of $\mathbf{Q}$.

Theorem 2—Under (C0)–(C2), if $\lambda_s > 2\mathrm{Re}\lambda_{s,2}$ ($s = 1,\ldots, r$), and $\gamma > 2\mathrm{Re}\gamma_2$, then for the design in (1) and (2) we have

$$\sqrt{n}\left(\frac{\mathbf{Y}_n}{n}-\gamma\mathbf{u},\frac{\mathbf{n}(n)}{n}-\mathbf{u}\right)\xrightarrow{d}\mathcal{N}(\mathbf{0},\mathbf{\Omega}_{Q,0}),\ \ \sqrt{n}\left(\frac{\mathbf{X}_n}{n}-\mathbf{u}\odot\lambda\odot\mathbf{V},\frac{\mathbf{N}(n)}{n}-\mathbf{u}\odot\mathbf{V}\right)\xrightarrow{d}\mathcal{N}(\mathbf{0},\mathbf{\Omega}_0),$$

where $\mathbf{\Omega}_{Q,0}$ is $\mathbf{\Omega}$ in Theorem 4 of YC, with $(\mathbf{M},\lambda_1,\mathbf{v}_1,\mathbf{a})$ there replaced by $(\mathbf{Q},\gamma,\mathbf{u},\mathbf{1})$ here; and set $\mathbf{\Lambda}_3^{\dagger}=\mathbf{\Lambda}_{23}^{\dagger}=\mathbf{\Lambda}_3^{\sharp}=\mathbf{\Lambda}_{23}^{\sharp}=\mathbf{\Lambda}_3^{\diamond}=\mathbf{\Lambda}_{23}^{\diamond}=\mathbf{\Lambda}_{32}^{\diamond}=\mathbf{0}$, $\mathbf{\Omega}_0$ is partitioned as $(\mathbf{\Omega}_{0,ij})_{1\le i,j\le 2}$ with

$$\mathbf{\Omega}_{0,11}=\sum_{s=1}^r u_s[(\mathbf{M}^{(s)})'\mathbf{\Lambda}^{\dagger}_{s,1}\mathbf{M}^{(s)}+\mathbf{\Lambda}^{\dagger}_{s,2}]+\boldsymbol{\lambda}\mathbf{V})'\odot\{\mathbf{\Lambda}_1^{\dagger}+(\mathbf{I}-\mathbf{u}'\mathbf{1})\mathbf{\Lambda}_2^{\sharp}(\mathbf{I}-\mathbf{1}'\mathbf{u})\}\odot(\boldsymbol{\lambda}\mathbf{V})+\sum_{1\le i\ne j\le r}[u_i\lambda_j(\mathbf{M}^{(i)})'\rho_{i,j}\mathbf{v}_j+u_i\lambda_j\mathbf{v}_j'\rho_{i,j}'\mathbf{M}^{(i)}]$$
$$\mathbf{\Omega}_{0,22}=\sum_{s=1}^r u_s[\mathbf{\Lambda}^{\dagger}_{s,1}+(\mathbf{I}-\mathbf{v}_s'\mathbf{1})\mathbf{\Lambda}^{\sharp}_{s,2}(\mathbf{I}-\mathbf{1}'\mathbf{v}_s)]+\mathbf{V}'\odot\{\mathbf{\Lambda}_1^{\dagger}+(\mathbf{I}-\mathbf{u}'\mathbf{1})\mathbf{\Lambda}_2^{\sharp}(\mathbf{I}-\mathbf{1}'\mathbf{u})\}\odot\mathbf{V}+\sum_{1\le i\ne j\le r}[u_i\rho_{i,j}\mathbf{v}_j+u_i\mathbf{v}_j'\rho_{i,j}']$$
$$\mathbf{\Omega}_{0,12}=\sum_{s=1}^r u_s[(\mathbf{M}^{(s)})'\mathbf{\Lambda}^{\dagger}_{s,1}+\mathbf{\Lambda}^{\diamond}_{s,2}(\mathbf{I}-\mathbf{1}'\mathbf{v}_s)]+(\boldsymbol{\lambda}\mathbf{V})'\odot\{\mathbf{\Lambda}_1^{\dagger}+(\mathbf{I}-\mathbf{u}'\mathbf{1})\mathbf{\Lambda}_2^{\sharp}(\mathbf{I}-\mathbf{1}'\mathbf{u})\}\odot\mathbf{V}+\sum_{1\le i\ne j\le r}[u_i(\mathbf{M}^{(i)})'\rho_{i,j}\mathbf{v}_j+\lambda_j u_i\mathbf{v}_i'\rho_{i,j}']$$

Remark: Note that if $r = 1$, then $\mathbf{\Lambda}_1^{\dagger}=\{\mathbf{u}\}-\mathbf{u}'\mathbf{u}=1-1'1=0$, $\mathbf{I}-\mathbf{u}'\mathbf{1} = 1 - 1'1 = 0$, and hence $(\boldsymbol{\lambda}\mathbf{V})'\odot\{\mathbf{\Lambda}_1^{\dagger}+(\mathbf{I}-\mathbf{u}'\mathbf{1})\Lambda_2^{\sharp}(\mathbf{I}-\mathbf{1}'\mathbf{u})\}\odot(\boldsymbol{\lambda}\mathbf{V})=\mathbf{0}$ and the last term in $\mathbf{\Omega}_0$ disappears. Thus, the effect of the urn $\mathbf{Q}$ is gone, and we go back to the case of YC.

As in Janson (2004) and Zhang et al. (2006), under regularity conditions, our design follows the functional central limit theorem for urn models.

3.2. Adaptive compound GPU design

The compound GPU design above is dependent on the unknown $\mathbf{F}$, and therefore it is not directly applicable in practice. A common method is to plug in its empirical estimate $\mathbf{F}_n = (F_{n,1},\ldots, F_{n,k})$ in the design, so that (1) and (2) are modified as

$$\theta_{n,sj}=\frac{q_{sj}(\mu_{n-1})\eta_{sj}}{\delta_{sj}},\ \ (j=1,\ldots,r),\ \ \theta_n=(\theta_{n,sj})_{r\times r} \tag{3}$$

$$\zeta^{(s)}_{n,ij}=\zeta^{(s)}_{n,ij}(\mu_{n-1},\xi^{(s)}_{ij})=\frac{m^{(s)}_{ij}(\mu_{n-1})\xi^{(s)}_{ij}}{v^{(s)}_{ij}},\ \ (j=1,\ldots,k),\ \ \boldsymbol{\zeta}^{(s)}_n=(\zeta^{(s)}_{n,ij})_{k\times k}. \tag{4}$$

Let $\mathcal{F}_m=\sigma\{\mathbf{X}^{(s)}_i,\mathbf{I}_i;I^{(s)}_i,\mathbf{r}^{(s)}_i:1\le s\le r;1\le i\le m\}$, where $\mathbf{I}_i = (I_{i1},\ldots,I_{ik})$. We impose the following assumptions:

(C3) $\zeta^{(s)}_{ij}(\cdot,\cdot)$ is continuous at $(\boldsymbol{\mu},\mathbf{y})$ for any $\mathbf{y}$ and there is an $r > 2$ such that

$$E\|\boldsymbol{\zeta}^{(s)}_1\|^r<\infty\ \text{ and } \sup_m E(\|\boldsymbol{\zeta}_{m,s}\|^r|\mathcal{F}_{m-1})<\infty,\ \ \text{a.s.}(s=1,\ldots r)$$

(C4) $\mathbf{M}^{(s)}(\boldsymbol{\mu}) > \mathbf{0}$ is bounded, differentiable, and there is a $\delta > 0$ such that for $s = 1,\ldots, r$,

$$\mathbf{M}^{(s)}(\mathbf{y})-\mathbf{M}^{(s)}(\mu)=\sum_{j=1}^{k}\frac{\partial\mathbf{M}^{s}(\mathbf{y})}{\partial y_j}\Big|_{\mathbf{y}=\mu}\|y_j-\mu_j\|+O(\|\mathbf{y}-\mu\|^{1+\delta}),\ \ \text{as } \|\mathbf{y}-\mu\|\to 0,$$

and

$$\mathbf{Q}(\mathbf{y})-\mathbf{Q}(\mu)=\sum_{j=1}^{k}\frac{\partial\mathbf{Q}(\mathbf{y})}{\partial y_j}\Big|_{\mathbf{y}=\mu}\|y_j-\mu_j\|+O(\|\mathbf{y}-\mu\|^{1+\delta}),\ \ \text{as } \|\mathbf{y}-\mu\|\to 0,$$

Let $S(\mathbf{F})$ be the support of $\mathbf{F}$.

Theorem 3—Assume that Assumptions (C0), (C1)–(C3). Then for the design given in (3) and (4) we have

- **i.** $\boldsymbol{\mu}_n \to \boldsymbol{\mu}$ *a.s.*, $\sup_{\mathbf{x}\in S(\mathbf{F})}\|\mathbf{F}_n(\mathbf{x}) - \mathbf{F}(\mathbf{x})\| \to 0$ *a.s.*
- **ii.** $(\mathbf{Y}_n/n,\mathbf{n}(n)/n,\mathbf{N}^{(s)}(n)/n^{(s)}) \to (\gamma\mathbf{u},\mathbf{u},\mathbf{v}_s)$ *a.s.* $(s = 1,\ldots,r)$,
- **iii.** $(\mathbf{X}_n/n,\mathbf{N}(n)/n) \to (\mathbf{u} \odot \boldsymbol{\lambda} \odot \mathbf{V},\mathbf{u} \odot \mathbf{V})$ *a.s.*

In order to state the asymptotic normality in the adaptive case, we need to introduce more notations. Define $\mathbf{t}_{s,j} = \mathbf{v}_s\partial\mathbf{M}(\mathbf{y})/\partial y_j|_{\mathbf{y}=\boldsymbol{\mu}}$, $\mathbf{T}^{(s)}=(\mathbf{t}'_{s,1}/v_{s,1},\ldots,\mathbf{t}'_{s,k}/v_{sk})'$, and $\mathbf{H}$ as before.

Define, for $s = 1,\ldots, r$,

$$\begin{aligned}
\mathbf{\Lambda}^{\dagger}_{s,3}&=\int_0^1\left[\int_x^1\frac{1}{y}\left(\frac{1}{y}\right)^{\overline{\mathbf{M}^{(s)}}}dy\right]'(\mathbf{T}^{(s)})'\sum\nolimits_3^{(s)}\mathbf{T}^{(s)}\left[\int_x^1\frac{1}{y}\left(\frac{1}{y}\right)^{\overline{\mathbf{M}^{(s)}}}dy\right]dx,\\
\mathbf{\Lambda}^{\sharp}_{s,3}&=\int_0^1\left[\int_x^1\int_x^y\frac{1}{yu}\left(\frac{y}{u}\right)^{\overline{\mathbf{M}^{(s)}}}du\,dy\right]'(\mathbf{T}^{(s)})'\sum\nolimits_3^{(s)}\mathbf{T}^{(s)}\left[\int_x^1\int_x^y\frac{1}{yu}\left(\frac{y}{u}\right)^{\overline{\mathbf{M}^{(s)}}}du\,dy\right]dx,\\
\mathbf{\Lambda}^{\diamond}_{s,3}&=\int_0^1\int_x^1\left[\frac{1}{y}\left(\frac{1}{y}\right)^{\overline{\mathbf{M}^{(s)}}}dy\right]'(\mathbf{T}^{(s)})'\sum\nolimits_3^{(s)}\mathbf{T}^{(s)}\int_x^1\int_x^y\frac{1}{yu}\left(\frac{y}{u}\right)^{\overline{\mathbf{M}^{(s)}}}du\,dy\,dx,\\
\mathbf{\Lambda}^{\sharp}_{s,4}&=\int_0^1\left[\int_x^1\int_x^y\frac{1}{yu}\left(\frac{y}{u}\right)^{\overline{\mathbf{M}^{(s)}}}du\,dy\right]'(\mathbf{T}^{(s)})'\sum\nolimits_4^{(s)}\mathbf{H}\left[\int_x^1\int_x^y\frac{1}{yu}\left(\frac{y}{u}\right)^{\overline{\mathbf{Q}}}du\,dy\right]dx,\\
\mathbf{\Lambda}^{\diamond}_{s,4}&=\int_0^1\int_x^1\left[\frac{1}{y}\left(\frac{1}{y}\right)^{\overline{\mathbf{M}^{(s)}}}dy\right]'(\mathbf{T}^{(s)})'\sum\nolimits_4^{(s)}\mathbf{H}\int_x^1\int_x^y\frac{1}{yu}\left(\frac{y}{u}\right)^{\overline{\mathbf{Q}}}du\,dy\,dx,\\
\mathbf{\Lambda}^{\dagger}_{s,23}&=\int_0^1\left(\frac{1}{x}\right)^{\overline{\mathbf{M}^{(s)}}'}\sum\nolimits_{23}^{(s)}\mathbf{T}^{(s)}\left[\int_x^1\frac{1}{y}\left(\frac{1}{y}\right)^{\overline{\mathbf{M}^{(s)}}}dy\right]dx,\\
\mathbf{\Lambda}^{\sharp}_{s,23}&=\int_0^1\left[\int_x^1\frac{1}{y}\left(\frac{y}{x}\right)^{\overline{\mathbf{M}^{(s)}}}dy\right]'(\mathbf{T}^{(s)})'\sum\nolimits_{23}^{(s)}\mathbf{T}^{(s)}\left[\int_x^1\frac{1}{y}\left(\frac{y}{x}\right)^{\overline{\mathbf{M}^{(s)}}}dy\right]dx,\\
\mathbf{\Lambda}^{\diamond}_{s,23}&=\int_0^1\left(\frac{1}{y}\right)^{\overline{\mathbf{M}^{(s)}}'}\sum\nolimits_{23}^{(s)}\mathbf{M}^{(s)}\int_x^1\frac{1}{yu}\int_x^y\left(\frac{y}{u}\right)^{\overline{\mathbf{M}^{(s)}}}du\,dy\,dx,
\end{aligned}$$

and

$$\mathbf{\Lambda}^{\diamond}_{s,32}=\int_0^1\left[\int_x^1\frac{1}{y}\left(\frac{1}{y}\right)^{\overline{\mathbf{M}^{(s)}}}dy\right]'\mathbf{H}'\left(\sum\nolimits_{23}^{(s)}\right)'\left[\int_x^1\frac{1}{y}\left(\frac{y}{x}\right)^{\overline{\mathbf{M}^{(s)}}}dy\right]dx.$$

Also, for convenience we recall the definition of $\mathbf{\Lambda}^{\dagger}_{3}, \mathbf{\Lambda}^{\sharp}_{3}, \mathbf{\Lambda}^{\diamond}_{3}, \mathbf{\Lambda}^{\dagger}_{23}, \mathbf{\Lambda}^{\sharp}_{23}, \mathbf{\Lambda}^{\diamond}_{23}$, and $\mathbf{\Lambda}^{\diamond}_{32}$.

$$\begin{aligned}
\mathbf{\Lambda}^{\dagger}_{3}&=\int_0^1\left[\int_x^1\frac{1}{y}\left(\frac{1}{y}\right)^{\overline{\mathbf{Q}}}dy\right]'\mathbf{H}'\sum\nolimits_3\mathbf{H}\left[\int_x^1\frac{1}{y}\left(\frac{1}{y}\right)^{\overline{\mathbf{Q}}}dy\right]dx,\\
\mathbf{\Lambda}^{\sharp}_{3}&=\int_0^1\left[\int_x^1\int_x^y\frac{1}{yx}\left(\frac{y}{u}\right)^{\overline{\mathbf{Q}}}dy\right]'\mathbf{H}'\sum\nolimits_3\mathbf{H}\left[\int_x^1\int_x^y\frac{1}{yu}\left(\frac{y}{u}\right)^{\overline{\mathbf{Q}}}du\,dy\right]dx,\\
\mathbf{\Lambda}^{\diamond}_{3}&=\int_0^1\int_x^1\left[\frac{1}{y}\left(\frac{y}{x}\right)^{\overline{\mathbf{Q}}}dy\right]'\mathbf{H}'\sum\nolimits_3\mathbf{H}\left[\int_x^1\int_x^y\frac{1}{yu}\left(\frac{y}{u}\right)^{\overline{\mathbf{Q}}}du\,dy\right]dx,\\
\mathbf{\Lambda}^{\dagger}_{23}&=\int_0^1\left(\frac{1}{x}\right)^{\overline{\mathbf{Q}}'}\sum\nolimits_{23}\mathbf{H}\left[\int_x^1\frac{1}{y}\left(\frac{1}{y}\right)^{\overline{\mathbf{Q}}}dy\right]dx,\\
\mathbf{\Lambda}^{\sharp}_{23}&=\int_0^1\left[\int_x^1\frac{1}{y}\left(\frac{y}{x}\right)^{\overline{\mathbf{Q}}}dy\right]'\mathbf{H}'\sum\nolimits_{23}\mathbf{H}\left[\int_x^1\frac{1}{y}\left(\frac{y}{x}\right)^{\overline{\mathbf{Q}}}dy\right]dx,\\
\mathbf{\Lambda}^{\diamond}_{23}&=\int_0^1\left(\frac{1}{y}\right)^{\overline{\mathbf{Q}}'}\sum\nolimits_{23}\mathbf{H}\int_x^1\frac{1}{yu}\int_x^y\left(\frac{y}{u}\right)^{\overline{\mathbf{Q}}}du\,dy\,dx,\\
\mathbf{\Lambda}^{\diamond}_{32}&=\int_0^1\left[\int_x^1\frac{1}{y}\left(\frac{1}{y}\right)^{\overline{\mathbf{Q}}}dy\right]'\mathbf{H}'\sum\nolimits_{23}'\left[\int_x^1\frac{1}{y}\left(\frac{y}{x}\right)^{\overline{\mathbf{Q}}}dy\right]dx.
\end{aligned}$$

Let $\mathbf{F}(\mathbf{x}) = (F_1(x_1), \ldots, F_k(x_k))'$ and $\mathbf{F}(\mathbf{x})(\mathbf{1} - \mathbf{F}(\mathbf{x})) = (F_1(x_1)(1 - F(x_1)),\ldots,F_k(x_k)(1 - F(x_k)))'$.

Theorem 4

i. Assuming (C0) and (C1), the following holds

a. $\sqrt{n}(\mu_n - \mu) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{\Omega}_\mu)$,

b. $\sqrt{n}(\mathbf{F}_n(\mathbf{x}) - \mathbf{F}(\mathbf{x})) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{\Omega}_F), \ \ \forall \mathbf{x} \in S(\mathbf{F})$,

where $\mathbf{\Omega}_\mu = \{\sum_{s=1}^r u_s \sigma^2/\mathbf{v}_s\}$ and $\mathbf{\Omega}_F = \{\sum_{s=1}^r u_s \mathbf{F}(\mathbf{x})(\mathbf{1} - \mathbf{F}(\mathbf{x}))/\mathbf{v}_s\}$.

ii. Assume (C0), (C1) and (C3), (C4), $\lambda_s > 2\mathrm{Re}\lambda_{s,2}$ ($s = 1,\ldots, r$), and $\gamma > 2\mathrm{Re}\gamma_2$. Then, for the design given in (3) and (4) the following holds

a. $\sqrt{n}\left(\frac{\mathbf{Y}_n}{n} - \gamma\mathbf{u}, \frac{\mathbf{n}(n)}{n} - \mathbf{u}\right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{\Omega}_Q)$,

b. $\sqrt{n}\left(\frac{\mathbf{X}_n}{n} - \mathbf{u} \odot \boldsymbol{\lambda} \odot \mathbf{V}, \frac{\mathbf{N}(n)}{n} - \mathbf{u} \odot \mathbf{V}\right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{\Omega})$,

where $\mathbf{\Omega}_Q$ is $\mathbf{\Omega}$ in Theorem 4 of YC, with ($\mathbf{M}$, λ_1, $\mathbf{v}_1$) there replaced by ($\mathbf{Q}$,γ,$\mathbf{u}$) here; $\mathbf{\Omega}$ is partitioned as $(\mathbf{\Omega}_{ij})_{1\le i,j\le 2}$ with

$$\begin{aligned}\mathbf{\Omega}_{11} = & \sum_{s=1}^r u_s[(\mathbf{M}^{(s)})' \mathbf{\Lambda}^\dagger_{s,1}\mathbf{M}^{(s)} + \mathbf{\Lambda}^\dagger_{s,2} + \mathbf{\Lambda}^\dagger_{s,3} + \mathbf{\Lambda}^\dagger_{s,23} + \mathbf{\Lambda}^{\dagger\,\prime}_{s,23}] \\ & + (\boldsymbol{\lambda}\mathbf{V})' \odot \{\mathbf{\Lambda}^\dagger_1 + (\mathbf{I} - \mathbf{u}'\mathbf{1})(\mathbf{\Lambda}^\sharp_2 + \mathbf{\Lambda}^\sharp_3)(\mathbf{I} - \mathbf{1}'\mathbf{u})\} \odot (\boldsymbol{\lambda}\mathbf{V}) \\ & + \sum_{1\le i\ne j\le r}[u_i\lambda_j\mathbf{v}'_j\rho'_{i,j}\mathbf{M}^{(i)} + u_i\lambda_j(\mathbf{M}^{(i)})'\rho_{i,j}\mathbf{v}_j + u_i\lambda_j\mathbf{\Lambda}^\diamond_{i,4}(\mathbf{I} - \mathbf{1}'\mathbf{u})_j\mathbf{v}_j + u_i\lambda_j\mathbf{v}'_j(\mathbf{I} - \mathbf{1}'\mathbf{u})'_j\mathbf{\Lambda}^{\diamond\,\prime}_{i,4}],\end{aligned}$$

$$\begin{aligned}\mathbf{\Omega}_{22} = & \sum_{s=1}^r u_s[\mathbf{\Lambda}^\dagger_{s,1} + (\mathbf{I} - \mathbf{v}'_s\mathbf{1})(\mathbf{\Lambda}^\sharp_{s,2} + \mathbf{\Lambda}^\sharp_{s,3} + \mathbf{\Lambda}^\sharp_{s,23} + \mathbf{\Lambda}^{\sharp\,\prime}_{s,23})(\mathbf{I} - \mathbf{1}'\mathbf{v}_s)] \\ & + \mathbf{V}' \odot \{\mathbf{\Lambda}^\dagger_1 + (\mathbf{1} - \mathbf{u}'\mathbf{1})(\mathbf{\Lambda}^\sharp_2 + \mathbf{\Lambda}^\sharp_3)(\mathbf{I} - \mathbf{1}'\mathbf{u})\} \odot \mathbf{V} \\ & + \sum_{1\le i\ne j\le r}[u_i\rho_{i,j}\mathbf{v}_j + u_j\mathbf{v}'_j\rho'_{i,j} + u_i(\mathbf{I} - \mathbf{1}'\mathbf{v}_i)'\mathbf{\Lambda}^\sharp_{i,4}(\mathbf{I} - \mathbf{1}'\mathbf{u})_j\mathbf{v}_j + u_i\mathbf{v}'_j(\mathbf{I} - \mathbf{1}'\mathbf{u})'_j\mathbf{\Lambda}^{\sharp\,\prime}_{i,4}(\mathbf{I} - \mathbf{1}'\mathbf{v}_i)],\end{aligned}$$

$$\begin{aligned}\mathbf{\Omega}_{12} = & \sum_{s=1}^r u_s[(\mathbf{M}^{(s)})'\mathbf{\Lambda}^\dagger_{s,1} + \mathbf{\Lambda}^\diamond_{s,2} + \mathbf{\Lambda}^\diamond_{s,3} + \mathbf{\Lambda}^\diamond_{s,23} + \mathbf{\Lambda}^\diamond_{s,32})(\mathbf{I} - \mathbf{1}'\mathbf{v}_s)] \\ & + \sum_{1\le i\ne j\le r}[u_i(\mathbf{M}^{(i)})'\rho_{i,j}\mathbf{v}_j + \lambda_i u_i\mathbf{v}'_i\rho'_{i,j} + u_i\mathbf{\Lambda}^\diamond_{i,4}(\mathbf{I} - \mathbf{1}'\mathbf{u})_j\mathbf{v}_j + \lambda_i u_i\mathbf{v}'_i((\mathbf{I} - \mathbf{1}'\mathbf{u})'_i\mathbf{\Lambda}^\sharp_{j,4})'(\mathbf{I} - \mathbf{1}'\mathbf{v}_j)] \\ & + (\boldsymbol{\lambda}\mathbf{V})' \odot \{\mathbf{\Lambda}^\dagger_1 + (\mathbf{I} - \mathbf{u}'\mathbf{1})(\mathbf{\Lambda}^\sharp_2 + \mathbf{\Lambda}^\sharp_3)(\mathbf{I} - \mathbf{1}'\mathbf{u})\} \odot \mathbf{V}.\end{aligned}$$

Here, $(\mathbf{I} - \mathbf{1}'\mathbf{u})_i$ is used to denote the i-th column of the matrix $(\mathbf{I} - \mathbf{1}'\mathbf{u})$.

The $\mathbf{\Omega}$ in Theorem 4(ii) is not easy to compute in general, but if we choose some special design, it can have a simple form as follows

<u>**Corollary:**</u> Let $\mathbf{u}(\cdot)$ and $\mathbf{v}_s(\cdot)$ be twice differentiable with $|\mathbf{u}(\cdot)| = 1$ and $|\mathbf{v}_s(\cdot)| = 1$. Set $\mathbf{Q}(\cdot)=\mathbf{1}_r'\mathbf{u}(\cdot)$ and $\mathbf{M}^{(s)}(\cdot)=\mathbf{1}_k'\mathbf{v}_s(\cdot)$. Assume that $\boldsymbol{\zeta}^{(s)}$ (s = 1,…, r) and $\boldsymbol{\theta}$ are constant matrices. Then, all the conditions of Theorem 4 are satisfied and $\boldsymbol{\Omega}$ in Theorem 4(ii) is simplified as

$$
\begin{aligned}
\boldsymbol{\Omega}_{11}=&2\sum_{s=1}^{r}u_s\boldsymbol{\Psi}_s\\
&+(\boldsymbol{\lambda}\mathbf{V})'\odot(\{\mathbf{u}\}\\
&-\mathbf{u}'\mathbf{u}\\
&+6(\mathbf{I}-\mathbf{1}'\mathbf{u})'\boldsymbol{\Phi}(\mathbf{I}-\mathbf{1}'\mathbf{u}))\odot(\boldsymbol{\lambda}\mathbf{V})\\
&+\sum_{1\le i\ne j\le r}[3u_i\lambda_i\mathbf{1}_{k\times k}\boldsymbol{\Gamma}_i\mathbf{1}_{r\times r}(\mathbf{I}-\mathbf{1}'\mathbf{u})_j\mathbf{v}_j+3u_i\lambda_j\mathbf{v}_j'(\mathbf{I}-\mathbf{1}'\mathbf{u})_j'(\mathbf{1}_{k\times k}\boldsymbol{\Gamma}_i\mathbf{1}_{r\times r})'+u_i\lambda_i(\mathbf{M}^{(i)})'\rho_{i,j}\mathbf{v}_j+u_i\lambda_j\mathbf{v}_j'\rho_{i,j}'\mathbf{M}^{(i)}],
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{\Omega}_{22}=&\sum_{s=1}^{r}u_s[\{\mathbf{v}_s\}-\mathbf{v}_s'\mathbf{v}_s+6(\mathbf{I}-\mathbf{v}_s'\mathbf{1})\boldsymbol{\Psi}_s(\mathbf{I}-\mathbf{1}'\mathbf{v}_s)]+\mathbf{V}'\odot(\{\mathbf{u}\}\\
&-\mathbf{u}'\mathbf{u}\\
&+6(\mathbf{I}-\mathbf{1}'\mathbf{u})'\boldsymbol{\Phi}(\mathbf{1}\\
&-\mathbf{1}'\mathbf{u}))\odot\mathbf{V}\\
&+\sum_{1\le i\ne j\le r}[6u_i(\mathbf{I}-\mathbf{1}'\mathbf{v}_i)'\mathbf{1}_{k\times k}\boldsymbol{\Gamma}_i\mathbf{1}_{r\times r}(\mathbf{I}-\mathbf{1}'\mathbf{u})_j\mathbf{v}_j+6u_i\mathbf{v}_j'(\mathbf{I}-\mathbf{1}'\mathbf{u})_j'(\mathbf{1}_{k\times k}\boldsymbol{\Gamma}_i\mathbf{1}_{r\times r})'(\mathbf{I}-\mathbf{1}'\mathbf{v}_i)+u_i\rho_{i,j}\mathbf{v}_j+u_i\mathbf{v}_j'\rho_{i,j}'],
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{\Omega}_{12}=&3\sum_{s=1}^{r}u_s[\boldsymbol{\Psi}_s(\mathbf{I}-\mathbf{1}'\mathbf{v}_s)]\\
&+(\boldsymbol{\lambda}\mathbf{V})'\odot(\{\mathbf{u}\}-\mathbf{u}'\mathbf{u}+6(\mathbf{I}-\mathbf{1}'\mathbf{u})'\boldsymbol{\Phi}(\mathbf{I}\\
&-\mathbf{1}'\mathbf{u}))\odot\mathbf{V}\\
&+\sum_{1\le i\ne j\le r}[3u_i\mathbf{1}_{k\times k}\boldsymbol{\Gamma}_i\mathbf{1}_{r\times r}(\mathbf{I}-\mathbf{1}'\mathbf{u})_j\mathbf{v}_j+6u_i\mathbf{v}_i'(\mathbf{I}-\mathbf{1}'\mathbf{u})_i'(\mathbf{1}_{k\times k}\boldsymbol{\Gamma}_i\mathbf{1}_{r\times r})'(\mathbf{I}-\mathbf{1}'\mathbf{v}_i)'+u_i(\mathbf{M}^{(i)})'\rho_{i,j}\mathbf{v}_j+\lambda_iu_i\mathbf{v}_i'\rho_{i,j}'],
\end{aligned}
$$

where $\boldsymbol{\Psi}_s = (\partial\mathbf{v}_s(\boldsymbol{\mu})/\partial\boldsymbol{\mu})'\{\boldsymbol{\sigma}^2/\mathbf{v}_s\}\partial\mathbf{v}_s(\boldsymbol{\mu})/\partial\boldsymbol{\mu}$, $\boldsymbol{\varphi} = (\partial\mathbf{u}(\boldsymbol{\mu})/\partial\boldsymbol{\mu})'\{\boldsymbol{\sigma}^2/\mathbf{u}\}\partial\mathbf{u}(\boldsymbol{\mu})/\partial\boldsymbol{\mu}$, and $\boldsymbol{\Upsilon}_s = (\partial\mathbf{v}_s(\boldsymbol{\mu})/\partial\boldsymbol{\mu})'(\boldsymbol{\sigma}^2\mathbf{v}_s)'\mathbf{e}_s\partial\mathbf{u}(\boldsymbol{\mu})/\partial\boldsymbol{\mu}$.

4. Application

Once the $\mathbf{v}_s$'s are obtained from the optimality criteria, the design matrices $\mathbf{M}^{(s)}$'s are constructed and the designs $\xi_{n,ij}^{(s)}$'s are easily obtained. The simplest such constructions are those given in the Corollary from the previous section; the general constructions for some commonly used designs in clinical trials can be found in YC with similar adding rules. Given the $G_s(\cdot)$'s, we can give an example to illustrate how to choose the $\mathbf{v}_s$'s so that the adaptive design is asymptotically optimal by the given joint criteria. In the following example, the $\mathbf{v}_s(\cdot)$'s (and $\mathbf{u}$) should be normalized and chosen differentiable around $\boldsymbol{\mu}$ to satisfy the required conditions.

4.1. An illustrative example

Now lets come back at the example briefed in the Introduction. The existing methods use different criteria separately. In this work, we propose to use them jointly. Let us consider these two commonly used criteria in clinical trials: the Neyman allocation (Melfi and Page, 1998) and the criterion in Rosenberger et al. (2001).

For two qualitative treatments, the Neyman allocation puts $N_1(n)$ patients to treatment I by maximizing the power for testing the difference $\Delta = p_1 - p_2$, where $\mathbf{p} = (p_1, p_2)$ are the success rates of the two treatments. Here $k = r = 2$, $f(\cdot) \equiv 1$, and $\boldsymbol{\mu} = \mathbf{p}$. This strategy leads to the allocation ratio

$$\frac{N_1(n)}{N_2(n)} \to \frac{v_{11}}{v_{12}} = \sqrt{\frac{p_1(1-p_1)}{p_2(1-p_2)}},$$

or $\mathbf{v}_1(\mathbf{p}) = (v_{11}(\mathbf{p}), v_{12}(\mathbf{p}))$, with

$$v_{11}(\mathbf{p}) = \frac{\sqrt{p_1(1-p_1)}}{\sqrt{p_1(1-p_1)} + \sqrt{p_2(1-p_2)}}, \quad v_{12}(\mathbf{p}) = \frac{\sqrt{p_2(1-p_2)}}{\sqrt{p_1(1-p_1)} + \sqrt{p_2(1-p_2)}}.$$

We have $|\mathbf{v}_1| = 1$ and the corresponding optimal $\mathbf{v}_1(\cdot) = (v_{11}(\cdot), v_{12}(\cdot))'$. The generating matrix $\mathbf{M}^{(1)}(\cdot)$ can be chosen as in the Corollary, $\mathbf{M}^{(1)}(\cdot) = \mathbf{1}_k' \mathbf{v}_1(\cdot)$. Take $\lambda_1 = 1$, and the $\xi_{n,ij}^{(1)}$'s be constants, the corresponding adding rule is $\boldsymbol{\varsigma}_{ni}^{(1)} = \mathbf{v}_1(\mathbf{p}_{n-1})\ (i=1,2)$. Note $\sigma_1^2 = p_1(1-p_1), \sigma_2^2 = p_2(1-p_2)$, and with the above choice of $\mathbf{M}_{(1)}$, it is easily verified that

$$\boldsymbol{\Psi}_1 = \frac{(1-2p_1)^2\sqrt{p_2(1-p_2)} + (1-2p_2)^2\sqrt{p_1-(1-p_1)}}{\left(\sqrt{p_1(1-p_1)} + \sqrt{p_2(1-p_2)}\right)^3} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

On the other hand, the criterion in Rosenberger et al. (2001) is to minimize the expected treatment failure and leads to

$$\frac{N_1(n)}{N_2(n)} \to \frac{v_{21}}{v_{22}} = \sqrt{\frac{p_1}{p_2}} \text{ or } v_{22}(\mathbf{p}) = \left(1 + \sqrt{\frac{p_1}{p_2}}\right)^{-1},$$

so that we can take the targeting $\mathbf{v}_2(\cdot)$ as

$$v_{21}(\mathbf{p}) = \frac{\sqrt{p_1}}{\sqrt{p_1} + \sqrt{p_2}}, \quad v_{22}(\mathbf{p}) = \frac{\sqrt{p_2}}{\sqrt{p_1} + \sqrt{p_2}}.$$

In practice, each trial is carefully selected but not perfect, so p_1 and p_2 are not close to zero and one, $p_1 \neq p_2$ and $\mathbf{v}_1$ and $\mathbf{v}_2$ can be very different. Apparently, a compromised criterion

will have the advantage of the two and avoid either extreme. The $f(\cdot)$,$\boldsymbol{\mu}$, the generating matrix and the adding rule are the same as above. It can be verified that

$$\boldsymbol{\Psi}_2=\frac{(1-p_1)p_2^{3/2}+(1-p_2)p_1^{3/2}}{\sqrt{p_1p_2}(\sqrt{p_1}+\sqrt{p_2})^3}\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}.$$

To join these two criteria, we choose $\mathbf{u}=(1-p_1+p_2, p_1-p_2)$ (suppose $p_1>p_2$). This design favors the first criterion when the difference p_1-p_2 is small, and the second otherwise. With the joint criteria, take $\boldsymbol{\lambda}=(1,1)$, $\mathbf{v}_1=(v_{11}, v_{12})$, $\mathbf{v}_2=(v_{21}, v_{22})$, and $\mathbf{V}=\begin{pmatrix}v_{11} & v_{12}\\ v_{21} & v_{22}\end{pmatrix}$. Then, we can easily show that

$$\boldsymbol{\Phi}=\left(\frac{\sigma_1^2}{1-p_1+p_2}+\frac{\sigma_2^2}{p_1-p_2}\right)\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}.$$

$$\boldsymbol{\Gamma}_1=\left(\frac{\sigma_1^2(1-2p_1)\sqrt{p_2(1-p_2)}}{2(\sqrt{p_1(1-p_1)}+\sqrt{p_2(1-p_2)})^3}+\frac{\sigma_2^2(1-2p_2)\sqrt{p_1(1-p_1)}}{2(\sqrt{p_1(1-p_1)}+\sqrt{p_2(1-p_2)})^3}\right)\begin{pmatrix}1 & -1\\ 1 & -1\end{pmatrix},$$

and

$$\boldsymbol{\Gamma}_2=\left(\frac{\sigma_1^2\sqrt{p_2}}{2(\sqrt{p_1}+\sqrt{p_2})^3}+\frac{\sigma_2^2\sqrt{p_1}}{2(\sqrt{p_1}+\sqrt{p_2})^3}\right)\begin{pmatrix}-1 & 1\\ -1 & 1\end{pmatrix}.$$

By Theorem 4, we have

$$\left(\frac{\mathbf{X}_n}{n},\frac{\mathbf{N}(n)}{n}\right)\to(\mathbf{u}\odot\lambda\odot\mathbf{V},\mathbf{u}\odot\mathbf{V})\text{ a.s,}$$

and by the Corollary,

$$\sqrt{n}\left(\frac{\mathbf{X}_n}{n}-\mathbf{u}\odot\lambda\odot\mathbf{V},\frac{\mathbf{N}(n)}{n}-\mathbf{u}\odot\mathbf{V}\right)\xrightarrow{d}\mathcal{N}(\mathbf{0},\boldsymbol{\Omega}),$$

with $\boldsymbol{\Omega}=(\boldsymbol{\Omega}_{ij})_{1\le i,j\le 2}$, where the $\boldsymbol{\Omega}_{ij}$'s are given by the corollary. Thus, the new allocation proportions under the compound method is

$$\mathbf{u}\odot\mathbf{V}=\frac{1-p_1+p_2}{\sqrt{p_1(1-p_1)}+\sqrt{p_2(1-p_2)}}(\sqrt{p_1(1-p_1)},\ \sqrt{p_2(1-p_2)})+\frac{p_1-p_2}{\sqrt{p_1}+\sqrt{p_2}}(\sqrt{p_1},\ \sqrt{p_2}),$$

which is a compromise of the two allocations, and is more robust than either one of two criterion used along, in that it keeps much of the power in detecting the difference while maintain less treatment losses.

4.2. Discussion

The proposed compound model is a response-adaptive design, so it will be helpful to have a brief discussion of its relationship and compare it with some commonly used such designs in clinical trials.

Bai and Hu (2005) studied a general urn model via random generating matrix, which includes many types of existing urn models used in clinical trials. Our compound model, however, is not covered by this class of models, due to the two-stage design. Tymofyeyev et al. (2007) established two equivalent optimization criteria for treatments allocation and proposed the corresponding adaptive randomization procedure targets the criterion. Again our design is not directly comparable to their procedure, both from motivation and structures of the design. Hu et al. (2009) proposed a new family of response-adaptive design, which attain the Cramer–Rao lower bounds on the allocation variances for any allocation proportions. The play-the-winner type design by Ivanova (2003) also attains this bound. A general comparison of the allocation proportion asymptotic variance of the proposed compound design with that of the lower bound of Hu, Zhang and He seems mathematically intractable. But intuitively our model will not attain this bound, as it takes into consideration of several optimality criteria together and thus has more variability than a design using a single criterion. We regard this as a weak aspect of our design. However our model is motivated to combine several optimality criteria into one design, so as to achieve a balanced goal, not for minimal variance alone. How to reduce the asymptotic variance of the compound design toward the lower bound will be a topic for our future study.

Acknowledgments

This research is partially supported by the National Center for Research Resources Grant No. 2G12RR003048 from the National Institutes of Health, USA.

References

Athreya KB, Karlin S. Limit theorems for the split times of branching processes. Journal of Mathematics and Mechanics. 1967; 17:257–277.

Bai ZD, Hu F. Asymptotics in randomized urn models. The Annals of Applied Probability. 2005; 1B: 914–940.

Eisele JR. The doubly adaptive biased coin design for sequential clinical trials. Journal of Statistical Planning and Inference. 1994; 38:249–262.

Eisele JR, Woodroofe M. Central limit theorems for doubly adaptive biased coin designs. Annals of Statistics. 1995; 23:234–254.

Gouet R. Strong convergence of proportions in a multicolor Pólya urn. Journal of Applied Probability. 1997; 34:426–435.

Hayre LS. Two-population sequential tests with three hypotheses. Biometrika. 1979; 66:465–474.

Hu F, Rosenberger WF. Optimality, variability, power: evaluating response-adaptive randomization procedures for treatment comparisons. Journal of the American Statistical Association. 2003; 98:671–678.

Hu, F.; Rosenberger, WF. The Theory of Response-adaptive Randomization in Clinical Trials. Wiley; Hoboken, NJ: 2006.

Hu F, Rosenberger WF, Zhang LX. Asymptotically best response-adaptive randomization procedures. Journal of Statistical Planning and Inference. 2006; 136:1911–1922.

Hu F, Zhang LX. Asymptotic properties of doubly adaptive biased coin designs for multitreatment clinical trials. Annals of Statistics. 2004; 32:268–301.

Hu F, Zhang LX, He X. Efficient randomized-adaptive designs. Annals of Statistics. 2009; 37:2543–2560.

Ivanova A. A play-the-winner-type urn design with reduced variability. Metrika. 2003; 58:1–13.

Janson S. Functional limit theorems for multitype branching processes and generalized Pólya urns. Stochastic Processes and Their Applications. 2004; 110:177–245.

Jennison, C.; Turnbull, BW. Group Sequential Methods with Applications to Clinical Trials. Chapman and Hall/CRC Press; Boca Raton, FL: 2000.

Melfi, V.; Page, C. Variability in adaptive designs for estimation of success probabilities. In: Flournoy, N.; Rosenberger, WF.; Wong, WK., editors. New Developments and Applications in Experimental Design. IMS; Hayward, CA: 1998. p. 106-114.

Rosenberger WF, Lachin JM. The use of response-adaptive designs in clinical trials. Controlled Clinical Trials. 1993; 14:471–484. [PubMed: 8119063]

Rosenberger WF, Stallard N, Ivanova A, Harper C, Ricks M. Optimal adaptive designs for binary response trials. Biometrics. 2001; 57:909–913. [PubMed: 11550944]

Tymofyeyev Y, Rosenberger WF, Hu F. Implementing optimal allocation in sequential binary response experiments. Journal of the American Statistical Association. 2007; 102:224–234.

Yuan A, Chai GX. Optimal adaptive generalized Pólya urn design for multi-arm clinical trials. Journal of Multivariate Analysis. 2008; 99:1–24.

Yuan A. Some practical issues of adaptive GPU design. Journal of Statistical Planning and Inference. 2008; 138:2953–2974.

Zhang L, Hu F, Cheung S. Asymptotic theorems of sequential estimation-adjusted urn models. Annals of Applied Probability. 2006; 16:340–369.

Appendix

Most of the techniques and notations below are from ZHC, with some new notations added. Define $\mathcal{M}_n=\sum_{m=1}^{n}\Delta\mathcal{M}_m$, $\mathcal{S}_n=\sum_{m=1}^{n}\Delta\mathcal{S}_m$, and $\mathcal{Q}_n=\sum_{m=1}^{n}\Delta\mathcal{Q}_m$, where $\Delta\mathcal{S}_m = \tilde{\mathbf{I}}_m\boldsymbol{\theta}_m - \tilde{E}[\tilde{\mathbf{I}}_m\boldsymbol{\theta}_m|\mathcal{F}_{m-1}]$, $\tilde{\mathbf{I}}_m=(I_m^{(1)},\ldots,I_m^{(r)})$, $\Delta\mathcal{M}_m = \tilde{\mathbf{I}}_m - \tilde{E}[\tilde{\mathbf{I}}_m|\mathcal{F}_{m-1}]$, and $\Delta\mathcal{Q}_m = \tilde{\mathbf{I}}_m\{\mathbf{f}(\tilde{\mathbf{r}}_m)-\tilde{\boldsymbol{\mu}}_m\}$. Here, $\mathbf{f}(\mathbf{r}_m) = (f(\mathbf{r}_m)|B_1, \ldots, f(\mathbf{r}_m)|B_r)$,$\boldsymbol{\mu}_m = (\mu_{m1}, \ldots, \mu_{mr})$, with $\mu_{ms} = E[f(\mathbf{r}_m)|\mathcal{F}_{m-1},B_s]$. We then have $\mathbf{n}- n\mathbf{u}=\mathbf{W}_n + o(n^{1/2})$ a.s., where

$$\mathbf{W}_n=\sum_{m=1}^{n}(\Delta\mathcal{M}_m+\Delta\mathcal{S}_m\mathcal{B}_{n,m}^{[2]}(\mathbf{I}-\mathbf{1}'\mathbf{u})+\Delta\mathcal{Q}_m\mathcal{R}\mathcal{B}_{n,m}^{[3]}(\mathbf{I}-\mathbf{1}'\mathbf{u})),$$

with $\overline{\mathcal{B}}_{n,m}=\prod_{j=m+1}^{n}(\mathbf{I}+\overline{\mathbf{Q}}/j)$, $\overline{\mathcal{B}}_{n,n} = \mathbf{I}$, $\overline{\mathcal{B}}_{n,m}^{[1]}=\sum_{j=m}^{n}\mathcal{B}_{n,j}/j$, $\mathcal{B}_{n,m}^{[2]}\sum_{j=m}^{n-1}\overline{\mathcal{B}}_{j,m}/j+1$, and $\mathcal{B}_{n,m}^{[3]}=\sum_{i=m}^{n-1}\sum_{j=m}^{i}\overline{\mathcal{B}}_{i,j}/(i+1)j$. Also, for $s=1,\ldots,r$, let $\mathbf{S}_n^{(s)}=\sum_{m=1}^{n}\Delta\mathbf{S}_m^{(s)}$, $\mathbf{m}_n^{(s)}=\sum_{m=1}^{n}\Delta\mathbf{m}_m^{(s)}$, and $\mathbf{Q}_n^{(s)}=\sum_{m=1}^{n}\Delta\mathbf{Q}_m^{(s)}$, where $\Delta\mathbf{S}_m^{(s)}=I_m^{(s)}\mathbf{I}_m\zeta_m^{(s)}-E[I_m^{(s)}\mathbf{I}_m\zeta_m^{(s)}|\mathcal{F}_{m-1}]$, $\Delta\mathbf{m}_m^{(s)}=I_m^{(s)}\mathbf{I}_m-E[I_m^{(s)}\mathbf{I}_m|\mathcal{F}_{m-1}]$, and $\Delta\mathbf{Q}_n^{(s)}=I_m^{(s)}\mathbf{I}_m\{\mathbf{f}(\mathbf{r}_m^{(s)})-\mu^{(s)}\}$. Here, $\mathbf{f}(\mathbf{r}_m)|B_s = (f(\mathbf{r}_m)|(A_1,B_s), \ldots, f(\mathbf{r}_m)|(A_k,B_s)$, $\mu^{(s)}=(\mu_1^{(s)},\ldots,\mu_k^{(s)})$ and note that in our case $\boldsymbol{\mu}^{(s)} = \boldsymbol{\mu}$. Then $\mathbf{S}_n^{(s)}$, $\mathbf{m}_n^{(s)}$, and $\mathbf{Q}_n^{(s)}$ are martingales. As in (5.19) and (5.20) in ZHC, or in the proof of Theorem 4 in YC, for each fixed s, we have

$$\mathbf{X}_n^{(s)}-n^{(s)}\lambda_s\mathbf{v}_s=\mathbf{U}_n^{(s)}+\mathrm{o}((n^{(s)})^{1/2})=\mathbf{U}_n^{(s)}+\mathrm{o}(n^{1/2})\ \text{(a.s.)}.$$

Similarly,

$$\mathbf{N}_n^{(s)} - n^{(s)}\mathbf{v}_s = \mathbf{V}_n^{(s)} + \mathrm{o}((n^{(s)})^{1/2}) = \mathbf{V}_n^{(s)} + \mathrm{o}(n^{1/2}) \ \text{(a.s.)}.$$

Here,

$\mathbf{U}_n^{(s)} = \sum_{m=1}^{n^{(s)}} (\Delta \mathbf{S}_m^{(s)} \overline{\mathbf{B}}_{n,m}^{(s)} + \Delta \mathbf{Q}_m^{(s)} \mathbf{T}^{(s)} \mathbf{B}_{n,m}^{(s),1}), \ \mathbf{V}_n^{(s)} = \sum_{m=1}^{n^{(s)}} (\Delta \mathbf{m}_m^{(s)} + \Delta \mathbf{S}_m^{(s)} \mathbf{B}_{n,m}^{(s),2} (\mathbf{I} - \mathbf{1}' \mathbf{v}_s) + \Delta \mathbf{Q}_m^{(s)} \mathbf{T}^{(s)} \mathbf{B}_{n,m}^{(s),3} (\mathbf{I} - \mathbf{1}' \mathbf{v}_s))$, with $\overline{\mathbf{B}}_{n,m}^{(s)} = \prod_{j=m+1}^{n^{(s)}} (\mathbf{I} + \overline{\mathbf{M}^{(s)}}/j), \ \overline{\mathbf{B}}_{n,n}^{(s)} = \mathbf{I}, \ \mathbf{B}_{n,n}^{(s),1} = \sum_{j=m}^{n^{(s)}} \overline{\mathbf{B}}_{n,j}^{(s)}/j, \ \overline{\mathbf{B}}_{n,m}^{(s),2} = \sum_{j=m}^{n^{(s)}-1} \overline{\mathbf{B}}_{j,m}^{(s)}/j+1$, and $\mathbf{B}_{n,m}^{(s),3} = \sum_{i=m}^{n^{(s)}-1} \sum_{j=m}^{i} \overline{\mathbf{B}}_{i,j}^{(s)}/(i+1)j$.

Proof of Theorem 1 and 3

Note that Theorem 1 is a special case of Theorem 3(ii). We only prove Theorem 3.

i. Note that in our case, $\boldsymbol{\mu}^{(s)} = \boldsymbol{\mu}$, $\mathbf{F}^{(s)} = \mathbf{F}(s = 1, \ldots, r)$, and $\sum_{s=1}^{r} u_s = 1$. The proofs follow the same lines as in that of Theorem 3(i) in YC, and using the first result in (ii).

ii. The first result can be obtained directly from Theorem 3 in YC. For simplicity, the proofs in ZHC and YC are done for $\lambda_s = 1$. However, these steps are valid for any value of λ_s and u_s, as shown in the following proofs. As to the second result, note that $\mathbf{X}_n - n\mathbf{u} \odot \mathbf{V} = \sum_{s=1}^{r} (\mathbf{X}_n^{(s)} - n u_s \mathbf{v}_s)$ and $\mathbf{N}(n) - n\mathbf{u} \odot \mathbf{V} = \sum_{s=1}^{r} (\mathbf{N}^{(s)}(n) - n u_s \mathbf{v}_s)$. Let us now fix $i(1 \le i \le k)$ and $s(1 \le s \le r)$. For fixed s, it follows from Theorem 3(ii) in YC that

$$\frac{n^{(s)}}{n} = \frac{Y_{m,s}}{n} = \frac{1}{n} \sum_{m=1}^{n} I_m^{(s)} \to u_s \ \text{a.s.}$$

Noting $X_{n,i}^{(s)}/n = (n^{(s)}/n)(X_{n,i}^{(s)}/n^{(s)})$ and using the fact (as in ZHC or YC) that $X_{n,i}^{(s)}/n^{(s)} \to v_{si}$ a.s., it results that $X_{n,i}^{(s)}/n \to u_s v_{si}$a.s..

Hence

$$\frac{\mathbf{X}_n^{(s)}}{n} = \left(\frac{n^{(s)}}{n} \frac{X_{n.i}^{(s)}}{n^{(s)}}, \ \ i = 1, \ldots, k \right) \to (u_s v_{s1}, \ldots, u_s v_{sk}) \ \text{a.s.},$$

which, in turn, implies that

$$\frac{\mathbf{X}_n}{n} = \sum_{s=1}^{r} \frac{\mathbf{X}_n^{(s)}}{n} \to \left(\sum_{s=1}^{r} u_s v_{s1}, \ldots, \sum_{s=1}^{r} u_s v_{sk} \right) = \mathbf{u} \odot \mathbf{V} \ \text{a.s.}.$$

As to $\mathbf{N}(n)$, we write

$$\mathbf{N}^{(s)}(n)-nu_s\mathbf{v}_s=\mathbf{m}_n^{(s)}+\sum_{m=1}^{n-1}\frac{\mathbf{X}_m^{(s)}}{|\mathbf{X}_m^{(s)}|}\left(\frac{Y_{m,s}}{|Y_m|}-u_s\right)+u_s\sum_{m=0}^{n-1}\left(\frac{\mathbf{X}_m^{(s)}}{|\mathbf{X}_m^{(s)}|}-\mathbf{v}_s\right).$$

Thus,

$$\mathbf{N}^{(s)}(n)-nu_s\mathbf{v}_s=\mathrm{o}(n)+\sum_{m=1}^{n-1}\frac{\mathbf{X}_m^{(s)}}{|\mathbf{X}_m^{(s)}|}\mathrm{o}(1)+u_s\sum_{m=0}^{n-1}\mathrm{o}(1)=\mathrm{o}(n)\ \text{ a.s.}$$

Proof of Theorem 2 and 4

Note that Theorem 2 is a special case of the Theorem 4(ii). Indeed, the distribution function $\mathbf{F}=(F_1,\ldots,F_k)$ of $f(\mathbf{r}_n)|A_i$, $i=1,\ldots,k$ is unknown in the adaptive compound GPU design whereas it is known in the nonadaptive case. In the case of Theorem 2, we set $\mathbf{H}=\mathbf{0}$. Thus, results in Theorem 2 is obtained from those in Theorem 4 by setting $\boldsymbol{\Lambda}_3^{\dagger}=\boldsymbol{\Lambda}_{23}^{\dagger}=\boldsymbol{\Lambda}_3^{\sharp}=\boldsymbol{\Lambda}_{23}^{\sharp}=\boldsymbol{\Lambda}_3^{\diamond}=\boldsymbol{\Lambda}_{23}^{\diamond}=\boldsymbol{\Lambda}_{32}^{\diamond}=\mathbf{0}$, and $\boldsymbol{\Lambda}_{s,3}^{\dagger}=\boldsymbol{\Lambda}_{s,23}^{\dagger}=\boldsymbol{\Lambda}_{s,3}^{\sharp}=\boldsymbol{\Lambda}_{s,4}^{\sharp}=\boldsymbol{\Lambda}_{s,23}^{\sharp}=\boldsymbol{\Lambda}_{s,3}^{\diamond}=\boldsymbol{\Lambda}_{s,4}^{\diamond}=\boldsymbol{\Lambda}_{s,23}^{\diamond}=\boldsymbol{\Lambda}_{s,32}^{\diamond}=\mathbf{0}$, $(s=1,\ldots,r)$. We now prove Theorem 4.

i. Rewrite $\mu_n^{(s)}$ as $\mu_n^{(s)}=(\sum_{j=1}^{N_i^{(s)}(n)}Z_{ji}^{(s)}:i=1,\ldots,k)$, where, for each fixed (i,s), the $Z_{ji}^{(s)}$'s are the $f(\mathbf{r}_j)$'s which are allocated to the s-th urn and i-th treatment. We have $E(Z_{ji}^{(s)})=\mu_i$, $Var(Z_{ji}^{(s)})=\sigma_i^2$, and for $(i,j,s)\neq(p,q,r)$, $Z_{ji}^{(s)}$ and $Z_{pq}^{(r)}$ are independent. Note that

$$\sqrt{n}(\mu_n-\mu)=\left(\sum_{s=1}^{r}\sqrt{\frac{n^{(s)}}{n}}\sqrt{\frac{n^{(s)}}{N_i^{(s)}(n)}}\sqrt{N_i^{(s)}(n)}\frac{\sum_{j=1}^{N_i^{(s)}(n)}Z_{ij}^{(s)}-\mu_i}{N_i^{(s)}(n)}:i=1,\ldots,k\right).$$

By Lemma 1 in YC,

$$\left(\sqrt{N_i^{(s)}(n)}\frac{\sum_{j=1}^{N_i^{(s)}(n)}Z_{ij}^{(s)}-\mu_i}{N_i^{(s)}(n)}:i=1,\ldots,k\right)\xrightarrow{d}N(\mathbf{0},\{\sigma^2\}).$$

Also, by Theorem 3, $\sqrt{n^{(s)}/n}\to\sqrt{u_s}(\text{a.s.})$ and $\sqrt{n^{(s)}/N_i^{(s)}(n)}\to 1/\sqrt{v_{s,i}}(\text{a.s.})$. Thus, the desired result follows from Slutsky's Theorem. The proof of the weak convergence of $\sqrt{n}(\mathbf{F}_n(\mathbf{x})-\mathbf{F}(\mathbf{x}))$ is similar and omitted.

ii. The first result follows directly from ZHC or Theorem 4(ii) in YC. As far as the second result is concerned, define $\mathbf{U}_n=\sum_{s=1}^{r}\mathbf{U}_n^{(s)}$ and $\mathbf{V}_n=\sum_{s=1}^{r}\mathbf{V}_n^{(s)}$. It can be shown that the sequences $\mathbf{U}_n$ and $\mathbf{V}_n$ are martingales with respect to $\{\mathcal{F}_n\}$. We then have

$$\sqrt{n}\left(\frac{\mathbf{X}_n}{n}-\mathbf{u}\odot\boldsymbol{\lambda}\odot\mathbf{V}\right)=\frac{1}{\sqrt{n}}\sum_{s=1}^{r}(\mathbf{X}_n^{(s)}-n^{(s)}\lambda_s\mathbf{v}_s)+\sqrt{n}\sum_{s=1}^{r}\left(\frac{n^{(s)}}{n}-u_s\right)\lambda_s\mathbf{v}_s=\frac{1}{\sqrt{n}}\mathbf{U}_n+\sqrt{n}\left(\frac{\mathbf{n}}{n}-\mathbf{u}\right)\odot\boldsymbol{\lambda}\odot\mathbf{V}+\mathrm{o}(1),\quad \text{a.s.}$$

and

$$\begin{aligned}\sqrt{n}\left(\frac{\mathbf{N}_n}{n}-\mathbf{u}\odot\mathbf{V}\right)&=\sqrt{n}\left[\sum_{s=1}^{r}\frac{n^{(s)}}{n}\left(\frac{\mathbf{N}_n^{(s)}}{n^{(s)}}-\mathbf{v}_s\right)+\sum_{s=1}^{r}\left(\frac{n^{(s)}}{n}-u_s\right)\mathbf{v}_s\right]\\&=\frac{1}{\sqrt{n}}\mathbf{V}_n\\&\quad+\sqrt{n}\sum_{s=1}^{r}\left(\frac{n^{(s)}}{n}-u_s\right)\mathbf{v}_s\\&=\frac{1}{\sqrt{n}}\mathbf{V}_n\\&\quad+\sqrt{n}\left(\frac{\mathbf{n}}{n}-\mathbf{u}\right)\odot\mathbf{V}\\&\quad+\mathrm{o}(1)\ \ (\text{a.s.})\end{aligned}$$

For r=1, $\mathbf{n}/n-\mathbf{u} = n/n-1 = 0$. For r>1, $\mathbf{n}= (n^{(1)},\dots,n^{(r)})$. It follows from Theorem 3.3 in ZHC or Theorem 4 in YC that $\sqrt{n}(\frac{\mathbf{n}}{n}-\mathbf{u})\xrightarrow{d}\mathcal{N}(\mathbf{0},\check{\boldsymbol{\Omega}}_{22})$, where $\check{\boldsymbol{\Omega}}_{22}=\boldsymbol{\Lambda}_1^{\dagger}+(\mathbf{I}-\mathbf{u}'\mathbf{1})(\boldsymbol{\Lambda}_2^{\sharp}+\boldsymbol{\Lambda}_3^{\sharp}+\boldsymbol{\Lambda}_{23}^{\sharp}+(\boldsymbol{\Lambda}_{23}^{\sharp})')(\mathbf{I}-\mathbf{1}'\mathbf{u})$. Here, $\mathbf{I}$ is $r\times r$ identity matrix, $\mathbf{1}$ is r-dimensional vector of 1's. In fact, $\check{\boldsymbol{\Omega}}_{22}$ is the lower-right $r\times r$ block of $\boldsymbol{\Omega}_{Q,0}$. Now, for r>1, we get

$$\sqrt{n}\sum_{s=1}^{r}\left(\frac{n^{(s)}}{n}-u_s\right)\mathbf{v}_s=\sqrt{n}\left(\frac{\mathbf{n}}{n}-\mathbf{u}\right)\odot\mathbf{V}\xrightarrow{d}\mathcal{N}(\mathbf{0},\mathbf{V}'\odot\check{\boldsymbol{\Omega}}_{22}\odot\mathbf{V}),$$

and $\sqrt{n}\sum_{s=1}^{r}(\frac{n^{(s)}}{n}-u_s)\lambda_s\mathbf{v}_s=\sqrt{n}(\frac{\mathbf{n}}{n}-\mathbf{u})\boldsymbol{\lambda}\odot\mathbf{V}\xrightarrow{d}\mathcal{N}(\mathbf{0},(\boldsymbol{\lambda}\mathbf{V})'\odot\check{\boldsymbol{\Omega}}_{22}\odot(\boldsymbol{\lambda}\mathbf{V}))$.

As in ZHC or in YC, $\mathbf{U}_n$ satisfies Linderberg's condition for martingale. Thus, by martingale central limit theorem, $n^{-1/2}\mathbf{U}_n\xrightarrow{d}\mathcal{N}(\mathbf{0},\tilde{\boldsymbol{\Omega}}_{11})$, where $\tilde{\boldsymbol{\Omega}}_{11}=\lim_{n\to\infty}\frac{1}{n}\mathrm{Var}(\mathbf{U}_n|)$. Here, we refer the notation Var[$\mathbf{U}_n$|] to ZHC for the average successive conditional covariances. Here, by Theorem 4 in this paper, and Theorem 4 in YC,

$$\tilde{\boldsymbol{\Omega}}_{11}=\lim_{n\to\infty}\mathrm{Var}\left(\sum_{s=1}^{r}\frac{n^{(s)}}{n}\frac{U_n^{(s)}}{\sqrt{n^{(s)}}}|\right)=\lim_{n\to\infty}\sum_{s=1}^{r}\frac{n^{(s)}}{n}\mathrm{Var}\left(\frac{U_n^{(s)}}{\sqrt{n^{(s)}}}|\right)=\sum_{s=1}^{r}u_s[(\mathbf{M}^{(s)})'\boldsymbol{\Lambda}_{s,1}^{\dagger}\mathbf{M}^{(s)}+\boldsymbol{\Lambda}_{s,2}^{\dagger}+\boldsymbol{\Lambda}_{s,3}^{\dagger}+\boldsymbol{\Lambda}_{s,23}^{\dagger}+(\boldsymbol{\Lambda}_{s,23}^{\dagger})']\ \text{a.s.}$$

Similarly, $1/\sqrt{n}\mathbf{V}_n\xrightarrow{d}\mathcal{N}(\mathbf{0},\tilde{\boldsymbol{\Omega}}_{22})$, where

$$\tilde{\boldsymbol{\Omega}}_{22}=\lim_{n\to\infty}\frac{1}{n}\mathrm{Var}(\mathbf{V}_n|)=\sum_{s=1}^{r}u_s[\boldsymbol{\Lambda}^{\dagger}_{s,1}+(\mathbf{I}-\mathbf{v}_s'\mathbf{1})(\boldsymbol{\Lambda}^{\sharp}_{s,2}+\boldsymbol{\Lambda}^{\sharp}_{s,3}+\boldsymbol{\Lambda}^{\sharp}_{s,23}+(\boldsymbol{\Lambda}^{\sharp}_{s,23})')(\mathbf{I}-\mathbf{1}'\mathbf{v}_s)]\ \text{a.s.}$$

Finally, using $\sqrt{n}(\frac{\mathbf{n}}{n}-\mathbf{u})=\frac{\mathbf{w}_n}{\sqrt{n}}+\mathrm{o}(1)$, we then have

$$\sqrt{n}\left(\frac{\mathbf{X}_n}{n}-\mathbf{u}\odot\boldsymbol{\lambda}\odot\mathbf{V},\frac{\mathbf{N}_n}{n}-\mathbf{u}\odot\mathbf{V}\right)=\left(\frac{\mathbf{U}_n}{\sqrt{n}}+\sqrt{n}(\frac{\mathbf{n}}{n}-\mathbf{u})\odot\boldsymbol{\lambda}\odot\mathbf{V},\frac{\mathbf{V}_n}{\sqrt{n}}+\sqrt{n}(\frac{\mathbf{n}}{n}-\mathbf{u})\odot\mathbf{V}\right)+\mathrm{o}(1)\xrightarrow{d}\mathcal{N}(\mathbf{0},\boldsymbol{\Omega}),$$

where $\boldsymbol{\Omega}=(\boldsymbol{\Omega}_{ij})_{1\le i,j\le 2}$, with

$$\boldsymbol{\Omega}_{11}=\tilde{\boldsymbol{\Omega}}_{11}+(\boldsymbol{\lambda}\mathbf{V})'\check{\boldsymbol{\Omega}}_{22}\odot(\boldsymbol{\lambda}\mathbf{V})+\boldsymbol{\Gamma}_{11}+\boldsymbol{\Gamma}_{11}'.\ \text{Here, }\boldsymbol{\Gamma}_{11}=\lim_{n\to\infty}\frac{1}{n}\mathrm{Cov}[(\mathbf{U}_n,\mathbf{W}_n\odot\boldsymbol{\lambda}\odot\mathbf{V})|].$$
$$\boldsymbol{\Omega}_{22}=\tilde{\boldsymbol{\Omega}}_{22}+\mathbf{V}'\odot\check{\boldsymbol{\Omega}}_{22}\odot\mathbf{V}+\boldsymbol{\Gamma}_{22}+\boldsymbol{\Gamma}_{22}'.\ \text{Here, }\boldsymbol{\Gamma}_{22}=\lim_{n\to\infty}\frac{1}{n}\mathrm{Cov}[(\mathbf{V}_n,\mathbf{W}_n\odot\mathbf{V})|].$$
$$\boldsymbol{\Omega}_{12}=\tilde{\boldsymbol{\Omega}}_{12}+\boldsymbol{\Gamma}_{14}+\boldsymbol{\Gamma}_{23}+(\boldsymbol{\lambda}\mathbf{V})'\odot\check{\boldsymbol{\Omega}}_{22}\odot\mathbf{V}.$$

Here, by Theorem 4 in YC,

$$\tilde{\boldsymbol{\Omega}}_{12}=\lim_{n\to\infty}Cov[\mathbf{U}_n,\mathbf{V}_n)|]=\sum_{s=1}^{r}u_s[(\mathbf{M}^{(s)})'\boldsymbol{\Lambda}^{\dagger}_{s,1}+(\boldsymbol{\Lambda}^{\diamond}_{s,2}+\boldsymbol{\Lambda}^{\diamond}_{s,3}+\boldsymbol{\Lambda}^{\diamond}_{s,23}+\boldsymbol{\Lambda}^{\diamond}_{s,32})(\mathbf{I}-\mathbf{1}'\mathbf{v}_s)]\ \text{a.s.},$$
$$\boldsymbol{\Gamma}_{14}=\lim_{n\to\infty}\frac{1}{n}\mathrm{Cov}[(\mathbf{U}_n,\mathbf{W}_n\odot\mathbf{V})|]\ \text{a.s., and }\boldsymbol{\Gamma}_{23}=\lim_{n\to\infty}\frac{1}{n}\mathrm{Cov}[(\mathbf{W}_n\odot\boldsymbol{\lambda}\odot\mathbf{V},\mathbf{V}_n)|]\text{a.s..}$$

To compute $\boldsymbol{\Gamma}_{11}$,$\boldsymbol{\Gamma}_{22}$,$\boldsymbol{\Gamma}_{14}$,$\boldsymbol{\Gamma}_{23}$, we need to compute the following quantities. Fix s, s=1…, r. Using Theorem 3, we have

$$\begin{aligned}
&\mathrm{Var}[\Delta\mathbf{m}_m^{(s)}|\mathcal{F}_{m-1}]\\
&\quad=E[I_m^{(s)}\{\mathbf{I}_m\}|\mathcal{F}_{m-1}]\\
&\quad-E[I_m^{(s)}\mathbf{I}_m|\mathcal{F}_{m-1}]'E[I_m^{(s)}\mathbf{I}_m|\mathcal{F}_{m-1}]\\
&\quad=\{E[I_m^{(s)}\mathbf{I}_m|\mathcal{F}_{m-1}]\}\\
&\quad-E[I_m^{(s)}\mathbf{I}_m|\mathcal{F}_{m-1}]'E[I_m^{(s)}\mathbf{I}_m|\mathcal{F}_{m-1}]\\
&\quad=\frac{Y_{m-1,s}}{|\boldsymbol{Y}_{m-1}|}\cdot\left\{\frac{\mathbf{X}_{m-1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|}\right\}\\
&\quad-\frac{Y_{m-1,s}}{|\boldsymbol{Y}_{m-1}|}\cdot\frac{(\mathbf{X}_{m-1}^{(s)})'}{|\mathbf{X}_{m-1}^{(s)}|}\cdot\frac{Y_{m-1,s}}{|\boldsymbol{Y}_{m-1}|}\cdot\frac{\mathbf{X}_{m-1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|}\to u_s\{\mathbf{v}_s\}\\
&\quad-u_s^2\mathbf{v}_s'\mathbf{v}_s\ \ \text{a.s.}
\end{aligned}$$

$$
\begin{aligned}
&\mathrm{Var}[\Delta\mathbf{S}_m^{(s)}|\mathcal{F}_{m-1}]\\
&\quad=E[I_m^{(s)}(\boldsymbol{\zeta}_m^{(s)})^{'}\{\mathbf{I}_m\}\boldsymbol{\zeta}_m^{(s)}|\mathcal{F}_{m-1}]-E[I_m^{(s)}\mathbf{I}_m\boldsymbol{\zeta}_m^{(s)}|\mathcal{F}_{m-1}]^{'}E[I_m^{(s)}\mathbf{I}_m\boldsymbol{\zeta}_m^{(s)}|\mathcal{F}_{m-1}]=\frac{Y_{m-1,s}}{|\mathbf{Y}_{m-1}|}E[(\boldsymbol{\zeta}_m^{(s)}-\mathbf{M}_m^{(s)})^{'}\left\{\frac{\mathbf{X}_{m-1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|}\right\}(\boldsymbol{\zeta}_m^{(s)}-\mathbf{M}_m^{(s)})|\mathcal{F}_{m-1}]+\frac{Y}{|\mathbf{Y}}\\
&\quad-u_s^2(\mathbf{M}^{(s)})^{'}\mathbf{v}_s^{'}\mathbf{v}_s\mathbf{M}^{(s)}\ \text{a.s.}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}[\Delta\mathbf{Q}_m^{(s)}|\mathcal{F}_{m-1}]&=E[I_m^{(s)}\{I_{m1}(f(r_{m1})|(A_1,B_s)-\mu_1)^2,\ldots,I_{mk}(f(r_{mk})|(A_k,B_s)-\mu_k)^2\}|\mathcal{F}_{n-1}]\\
&=\left\{E[I_m^{(s)}I_{m1}(f(r_{m1})|A_1-\mu_1)^2|\mathcal{F}_{m-1}],\ldots,E[I_{mk}(f(r_{mk})|A_k-\mu_k)^2|\mathcal{F}_{m-1}]\right\}\\
&=\frac{Y_{m-1,s}}{|Y_{m-1}|}\left\{\sigma_1^2\frac{X_{m-1,1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|},\ldots\sigma_k^2\frac{X_{m-1,k}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|}\right\}\to u_s\sum\nolimits_3^{(s)}\ \text{a.s.}
\end{aligned}
$$

$$
\begin{aligned}
&\mathrm{Cov}[(\Delta\mathbf{S}_m^{(s)},\Delta\mathbf{m}_m^{(s)})|\mathcal{F}_{m-1}]\\
&\quad=E[I_m^{(s)}(\boldsymbol{\zeta}_m^{(s)})^{'}\{\mathbf{I}_m\}|\mathcal{F}_{m-1}]-E[I_m^{(s)}\mathbf{I}_m\boldsymbol{\zeta}_m^{(s)}|\mathcal{F}_{m-1}]^{'}E[I_m^{(s)}\mathbf{I}_m|\mathcal{F}_{m-1}]=\frac{Y_{m-1,s}}{|\mathbf{Y}_{m-1}|}(\mathbf{M}_m^{(s)})^{'}\left\{\frac{\mathbf{X}_{m-1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|}\right\}-\left(\frac{Y_{m-1,s}}{|\mathbf{Y}_{m-1}|}\right)^2(\mathbf{M}_m^{(s)})^{'}\frac{(\mathbf{X}_{m-1}^{(s)})^{'}}{|\mathbf{X}_{m-1}^{(s)}|}\cdot\frac{\mathbf{X}_{m-1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|}\to\\
&\quad-u_s^2(\mathbf{M}^{(s)})^{'}\mathbf{v}_s^{'}\mathbf{v}_s\ \text{a.s..}
\end{aligned}
$$

$$
\begin{aligned}
&\mathrm{Cov}[(\Delta\mathbf{m}_m^{(s)},\\
&\quad\Delta\mathbf{Q}_m^{(s)})|\mathcal{F}_{m-1}]\\
&\quad=\mathbf{0}\ \text{a.s.}\mathrm{Cov}[(\Delta\mathbf{S}_m^{(s)},\\
&\quad\Delta\mathbf{Q}_m^{(s)})|\mathcal{F}_{m-1}]=E[(I_m^{(s)}\mathbf{I}_n\boldsymbol{\zeta}_m^{(s)}-E[I_m^{(s)}\mathbf{I}_m\boldsymbol{\zeta}_m^{(s)}|\mathcal{F}_{m-1}])^{'}I_m^{(s)}\mathbf{I}_m\{f(\mathbf{r}_m)|B_s\\
&\quad-\mu\}|\mathcal{F}_{m-1}]\\
&\quad=\frac{Y_{m-1,s}}{|\mathbf{Y}_{m-1}|}E[(\boldsymbol{\zeta}_m^{(s)})^{'}\{\mathbf{I}_m\}\{f(\mathbf{r}_m)|B_s\\
&\quad-\mu\}|I_m^{(s)}\\
&\quad=1,\mathcal{F}_{m-1}].
\end{aligned}
$$

Using the fact that $\boldsymbol{\xi}^{(s)}$ is independent of all the other variables and that $f(\mathbf{r}_m)|B_s$ is independent of $\mathcal{F}_{m-1}$, we obtain:

$\mathrm{Cov}[(\Delta\mathbf{S}_m^{(s)},\Delta\mathbf{Q}_m^{(s)})|\mathcal{F}_{m-1}]=\mathbf{0}$ a.s., $E[(\Delta\mathbf{m}_m^{(s)})^{'}\Delta\mathcal{M}_m|\mathcal{F}_{m-1}]=E[I_m^{(s)}\mathbf{I}_m^{'}\tilde{\mathbf{I}}_m|\mathcal{F}_{m-1}]-E[I_m^{(s)}\mathbf{I}_m^{'}|\mathcal{F}_{m-1}]E[\tilde{\mathbf{I}}_m|\mathcal{F}_{m-1}]$. Note

$$
E[I_m^{(s)}\mathbf{I}_m^{'}\tilde{\mathbf{I}}_m|\mathcal{F}_{m-1}]=E[I_m^{(s)}\mathbf{I}_m^{'}|\mathcal{F}_{m-1}]\mathbf{e}_s=\frac{Y_{m-1,s}}{|Y_{m-1}|}\left(\frac{\mathbf{X}_{m-1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|}\right)^{'}\mathbf{e}_s.
$$

Thus, by Theorem 3.1 in ZHC, or Theorem 3 in YC,

$$E[(\Delta \mathbf{m}_m^{(s)})^{'} \Delta \mathcal{M}_m | \mathcal{F}_{m-1}] = \frac{Y_{m-1,s}}{|Y_{m-1}|} \left(\frac{\mathbf{X}_{m-1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|} \right)^{'} \mathbf{e}_s - \left(\frac{\mathbf{X}_{m-1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|} \right)^{'} \frac{Y_{m-1}}{|Y_{m-1}|} \to u_s \mathbf{v}_s^{'}(\mathbf{e}_s - \mathbf{u}) := u_s \sum^{(s)} \text{ a.s.},$$
$$E[(\Delta \mathbf{m}_m^{(s)})^{'} \Delta \mathcal{S}_m | \mathcal{F}_{m-1}] = E[I_m^{(s)} \mathbf{I}_m^{'} \tilde{\mathbf{I}}_m \theta_m | \mathcal{F}_{m-1}] - E[I_m^{(s)} \mathbf{I}_m^{'} | \mathcal{F}_{m-1}] E[\tilde{\mathbf{I}}_m \theta_m | \mathcal{F}_{m-1}].$$

But

$$E[I_m^{(s)} \mathbf{I}_m^{'} \tilde{\mathbf{I}}_m \theta_m | \mathcal{F}_{m-1}] = E[I_m^{(s)} \mathbf{I}_m^{'} \mathbf{e}_s \theta_m | \mathcal{F}_{m-1}] = E[I_m^{(s)} \mathbf{I}_m^{'} \theta_{m,s} | \mathcal{F}_{m-1}] = \frac{Y_{m-1,s}}{|\mathbf{Y}_{m-1}|} \left(\frac{\mathbf{X}_{m-1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|} \right)^{'} \tilde{\mathbf{q}}_{m,s}.$$

Here, $\tilde{\mathbf{q}}_{m,s}$ is the s-th row of the matrix $\mathbf{Q}_m$. Note that $\mathbf{Q}_m \xrightarrow{a.s.} \mathbf{Q}$ and denote $\tilde{\mathbf{q}}_s$ as the s-th row of $\mathbf{Q}$. Thus,

$$E[(\Delta \mathbf{m}_m^{(s)})^{'} \Delta \mathcal{S}_m | \mathcal{F}_{m-1}] = \frac{Y_{m-1,s}}{|Y_{m-1}|} \left(\frac{\mathbf{X}_{m-1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|} \right)^{'} \left(\tilde{\mathbf{q}}_{m,s} - \frac{Y_{m-1}}{|Y_{m-1}|} \mathbf{Q}_m \right) \to u_s \mathbf{v}_s^{'} (\tilde{\mathbf{q}}_s - \mathbf{u}\mathbf{Q}) \text{ a.s.},$$
$$E[(\Delta \mathbf{m}_m^{(s)})^{'} \Delta Q_m | \mathcal{F}_{m-1}] = E[I_m^{(s)} \mathbf{I}_m^{'} \tilde{\mathbf{I}}_m \{\mathbf{f}(\mathbf{r}_m) - \tilde{\mu}_m\} | \mathcal{F}_{m-1}] - E[I_m^{(s)} \mathbf{I}_m | \mathcal{F}_{m-1}] E[\tilde{\mathbf{I}}_m \{\mathbf{f}(\mathbf{r}_m) - \tilde{\mu}_m\} | \mathcal{F}_{m-1}] = \mathbf{0} \text{ a.s.},$$
$$E[(\Delta \mathbf{S}_m^{(s)})^{'} \Delta \mathcal{M}_m | \mathcal{F}_{m-1}] = E[I_m^{(s)} (\boldsymbol{\zeta}_m^{(s)})^{'} \mathbf{I}_m^{'} \tilde{\mathbf{I}}_m | \mathcal{F}_{m-1}] - E[I_m^{(s)} \boldsymbol{\zeta}_m^{(s)} \mathbf{I}_m^{'} | \mathcal{F}_{m-1}] E[\tilde{\mathbf{I}}_m | \mathcal{F}_{m-1}].$$

But

$$E[I_m^{(s)} (\boldsymbol{\zeta}_m^{(s)})^{'} \mathbf{I}_m^{'} \tilde{\mathbf{I}}_m | \mathcal{F}_{m-1}] = E[(\boldsymbol{\zeta}_m^{(s)})^{'} I_m^{(s)} \mathbf{I}_m^{'} | \mathcal{F}_{m-1}] \mathbf{e}_s = \frac{Y_{m-1,s}}{\mathbf{Y}_{m-1}} (\mathbf{M}_m^{(s)})^{'} \left(\frac{\mathbf{X}_{m-1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|} \right)^{'} \mathbf{e}_s.$$

Thus,

$$E[(\Delta \mathbf{S}_m^{(s)})^{'} \Delta \mathcal{M}_m | \mathcal{F}_{m-1}] = \frac{Y_{m-1,s}}{\mathbf{Y}_{m-1}} (\mathbf{M}_m^{(s)})^{'} \left(\frac{\mathbf{X}_{m-1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|} \right)^{'} \mathbf{e}_s - (\mathbf{M}_m^{(s)})^{'} \left(\frac{\mathbf{X}_{m-1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|} \right)^{'} \frac{Y_{m-1,s}}{|\mathbf{Y}_{m-1}|} \to u_s [(\mathbf{M}^{(s)})^{'} \mathbf{v}_s^{'} \mathbf{e}_s - (\mathbf{M}^{(s)})^{'} \mathbf{v}_s^{'} \mathbf{u}] \text{ a.s.},$$
$$E[(\Delta \mathbf{S}_m^{(s)})^{'} \Delta \mathcal{S}_m | \mathcal{F}_{m-1}] = E[I_m^{(s)} (\boldsymbol{\zeta}_m^{(s)})^{'} (\mathbf{I}_m)^{'} \tilde{\mathbf{I}}_m \theta_m | \mathcal{F}_{m-1}] - E[I_m^{(s)} (\boldsymbol{\zeta}_m^{(s)})^{'} (\mathbf{I}_m)^{'} | \mathcal{F}_{m-1}] E[\tilde{\mathbf{I}}_m \theta_m | \mathcal{F}_{m-1}].$$

But

$$E[I_m^{(s)} (\boldsymbol{\zeta}_m^{(s)})^{'} (\mathbf{I}_m)^{'} \tilde{\mathbf{I}}_m \theta_m | \mathcal{F}_{m-1}] = E[I_m^{(s)} (\boldsymbol{\zeta}_m^{(s)})^{'} \mathbf{I}_m^{'} \mathbf{e}_s \theta_m | \mathcal{F}_{m-1}] = E[I_m^{(s)} (\boldsymbol{\zeta}_m^{(s)})^{'} \mathbf{I}_m^{'} \theta_{m,s} | \mathcal{F}_{m-1}] = \frac{Y_{m-1,s}}{\mathbf{Y}_{m-1}} (\mathbf{M}_m^{(s)})^{'} \left(\frac{\mathbf{X}_{m-1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|} \right)^{'} \tilde{\mathbf{q}}_{m,s}.$$

Thus,

$$E[(\Delta\mathbf{S}_m^{(s)})'\Delta\mathcal{S}_m|\mathcal{F}_{m-1}]=\frac{Y_{m-1,s}}{\mathbf{Y}_{m-1}}(\mathbf{M}_m^{(s)})'\left(\frac{\mathbf{X}_{m-1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|}\right)'\tilde{\mathbf{q}}_{m,s}-\frac{Y_{m-1,s}}{\mathbf{Y}_{m-1}}(\mathbf{M}_m^{(s)})'\left(\frac{\mathbf{X}_{m-1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|}\right)'\frac{\mathbf{Y}_{m-1}}{|\mathbf{Y}_{m-1}|}\mathbf{Q}_m\to u_s[(\mathbf{M}^{(s)})'\mathbf{v}_s'\tilde{\mathbf{q}}_s-(\mathbf{M}^{(s)})'\mathbf{v}_s'\mathbf{u}\mathbf{Q}]\ \text{a.s.},$$
$$E[(\Delta\mathbf{S}_m^{(s)})'\Delta\mathcal{Q}_m|\mathcal{F}_{m-1}]=\mathbf{0}\ \text{a.s.},\ E[(\Delta\mathbf{Q}_m^{(s)})'\Delta\mathcal{M}_m|\mathcal{F}_{m-1}]=\mathbf{0}\ \text{a.s.}$$
$$E[(\Delta\mathbf{Q}_m^{(s)})'\Delta\mathcal{S}_m|\mathcal{F}_{m-1}]=\mathbf{0}\ \text{a.s.},\ E[(\Delta\mathbf{Q}_m^{(s)})'\Delta\mathcal{Q}_m|\mathcal{F}_{m-1}]=E[\{\mathbf{f}(\mathbf{r}_m)|B_s-\mu^{(s)}\}I_m^{(s)}(\mathbf{I}_m)'\tilde{\mathbf{I}}_{mm}\{\mathbf{f}(\mathbf{r}_m)-\tilde{\mu}_m\}|\mathcal{F}_{m-1}]$$
$$=E[\{\mathbf{f}(\mathbf{r}_m)|B_s-\mu^{(s)}\}I_m^{(s)}(\mathbf{I}_m)'\mathbf{e}_s\{\mathbf{f}(\mathbf{r}_m)-\tilde{\mu}_m\}|\mathcal{F}_{m-1}]=E[I_m^{(s)}\{\mathbf{f}(\mathbf{r}_m)|B_s-\mu^{(s)}\}(\mathbf{I}_m)'(f(\mathbf{r}_m)-\tilde{\mu}_{ms})\mathbf{e}_s|\mathcal{F}_{m-1}]$$
$$=P(B_s|\mathcal{F}_{m-1})E[\{\mathbf{f}(\mathbf{r}_m)|B_s-\mu^{(s)}\}(\mathbf{I}_m)'(\mathbf{f}(\mathbf{r}_m)-\tilde{\mu}_{ms})|\mathcal{F}_{m-1},B_s]]\mathbf{e}_s.$$

But for

$$\begin{aligned}j=1,\ldots,k,\ E[I_{mj}(f(\mathbf{r}_m)|(B_s,&\\ &A_j)-\mu_j)(f(\mathbf{r}_m)|B_s\\ &-\tilde{\mu}_{m,s})|\mathcal{F}_{m-1},\\ &B_s]]\\ &=P(A_j|\mathcal{F}_{m-1},B_s)E[(f(\mathbf{r}_m)|B_s-\tilde{\mu}_{m,s})(f(\mathbf{r}_m)|(B_s,A_j)-\mu_j)|\mathcal{F}_{m-1},B_s],A_j].\end{aligned}$$

Note that $\tilde{\mu}_{m,s}=\sum_{i=1}^{k}\mu_iP(A_j|\mathcal{F}_{m-1},B_s)]$ and $f(\mathbf{r}_m)|B_s-\tilde{\mu}_{m,s}=\sum_{i=1}^{k}[f(\mathbf{r}_m)|B_sI_{mi}-\mu_iP(A_j|\mathcal{F}_{m-1},B_s)]$. We can then write

$$\begin{aligned}&E[I_{mj}(f(\mathbf{r}_m)|(B_s,A_j)-\mu_j)(f(\mathbf{r}_m)|B_s-\tilde{\mu}_{m,s})|\mathcal{F}_{m-1},B_s\\ &=\frac{X_{m-1,j}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|}\left(\sum_{i=1}^{k}E[I_{mi}f(\mathbf{r}_m)|B_sf(\mathbf{r}_m)|A_j-\mu_if(\mathbf{r}_m)|A_jP(A_j|\mathcal{F}_{m-1},B_s)-I_{mi}\mu_jf(\mathbf{r}_m)|B_s+\mu_j\mu\right.\\ &=\frac{X_{m-1,j}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|}(E[f^2(\mathbf{r}_m)|A_j|\mathcal{F}_{m-1},A_j,B_s]-\mu_iE[f(\mathbf{r}_m)|A_jP(A_j|\mathcal{F}_{m-1},B_s)]|-\mathcal{F}_{m-1},A_j,B_s]-\mu_jE[f(\mathbf{r}_m)|A_j|\mathcal{F}_{m-1},A_j,B_s]+\mu_j^2E[P(A_j|\mathcal{F}_{m-1},B_s)|\mathcal{F}\end{aligned}$$

Thus,

$$E[(\Delta\mathbf{Q}_m^{(s)})'\Delta\mathcal{Q}_m|\mathcal{F}_{m-1}]=\frac{Y_{m-1,s}}{\mathbf{Y}_{m-1}}\left(\sigma_1^2\frac{X_{m-1,1}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|},\ldots,\sigma_k^2\frac{X_{m-1,k}^{(s)}}{|\mathbf{X}_{m-1}^{(s)}|}\right)'\mathbf{e}_s\to u_s((\sigma_1^2v_{s1},\ldots,\sigma_k^2v_{sk})'\mathbf{e}_s:=u_s\sum\nolimits_4^{(s)}\ \text{a.s.}$$

Now, let us write $\mathbf{U}_n$, $\mathbf{V}_n$, and $\mathbf{W}_n$ as follows

$$\mathbf{U}_n=\mathbf{U}_n^{[1]}+\mathbf{U}_n^{[2]}+\mathbf{U}_n^{[3]},\ \mathbf{V}_n=\mathbf{V}_n^{[1]}+\mathbf{V}_n^{[2]}+\mathbf{V}_n^{[3]},\ \text{ and }\ \mathbf{W}_n=\mathbf{W}_n^{[1]}+\mathbf{W}_n^{[2]}+\mathbf{W}_n^{[3]},$$

where

$$\mathbf{U}_n^{[1]}=\sum_{s=1}^{r}\sum_{m=1}^{n}\Delta\mathbf{m}_m^{(s)}\overline{\mathbf{B}}_{n,m}^{(s)}\mathbf{M}^{(s)},\ \mathbf{U}_n^{[2]}=\sum_{s=1}^{r}\sum_{m=1}^{n}(\Delta\mathbf{S}_m^{(s)}-\Delta\mathbf{m}_m^{(s)}\mathbf{M}^{(s)})\overline{\mathbf{B}}_{n,m}^{(s)},$$

$$\mathbf{U}_n^{[3]}=\sum_{s=1}^{r}\sum_{m=1}^{n}\Delta\mathbf{Q}_m^{(s)}\mathbf{T}^{(s)}\mathbf{B}_{n,m}^{(s),1},\ \mathbf{V}_n^{[1]}=\sum_{s=1}^{r}\sum_{m=1}^{n}\Delta\mathbf{m}_m^{(s)}\overline{\mathbf{B}}_{n,m}^{(s)},$$

$$\mathbf{V}_n^{[2]}=\sum_{s=1}^{r}\sum_{m=1}^{n}(\Delta\mathbf{S}_m^{(s)}-\Delta\mathbf{m}_m^{(s)}\mathbf{M}^{(s)})\mathbf{B}_{n,m}^{(s),2}(\mathbf{I}-\mathbf{1}'\mathbf{v}_s),$$

$$\mathbf{V}_n^{[3]}=\sum_{s=1}^{r}\sum_{m=1}^{n}\Delta\mathbf{Q}_m^{(s)}\mathbf{T}^{(s)}\mathbf{B}_{n,m}^{(s),3}(\mathbf{I}-\mathbf{1}'\mathbf{v}_s),\quad \mathbf{W}_n^{[1]}=\sum_{m=1}^{n}\Delta\mathcal{M}_m\overline{\mathcal{B}}_{n,m},$$

$$\mathbf{W}_n^{[2]}=\sum_{m=1}^{n}(\Delta\mathcal{S}_m-\Delta\mathcal{M}_m\mathbf{Q})\mathcal{B}_{n,m}^{[2]}(\mathbf{I}-\mathbf{1}'\mathbf{u}),\ \text{ and }\ \mathbf{W}_n^{[3]}=\sum_{m=1}^{n}\Delta\mathcal{Q}_m\mathcal{B}_{n,m}^{[3]}(\mathbf{I}-\mathbf{1}'\mathbf{u}).$$

We can now derive

$$\begin{aligned}\boldsymbol{\Gamma}_{11}=\lim_{n\to\infty}n^{-1}\sum_{m=1}^{n}\mathrm{Cov}[(\mathbf{U}_m,\mathbf{W}_n\odot\boldsymbol{\lambda}\odot\mathbf{V})|]=\lim_{n\to\infty}n^{-1}\sum_{m=1}^{n}\{&\mathrm{Cov}[(\mathbf{U}_m^{[1]},\mathbf{W}_m^{[1]}\odot\boldsymbol{\lambda}\odot\mathbf{V})|]+\mathrm{Cov}[(\mathbf{U}_m^{[1]},\mathbf{W}_m^{[2]}\odot\boldsymbol{\lambda}\odot\mathbf{V})|]\\ +\mathrm{Cov}[(\mathbf{U}_m^{[1]},\mathbf{W}_m^{[3]}\odot\boldsymbol{\lambda}\odot\mathbf{V})|]+\mathrm{Cov}[(\mathbf{U}_m^{[2]},\mathbf{W}_m^{[1]}\odot\boldsymbol{\lambda}\odot\mathbf{V})|]&+\mathrm{Cov}[(\mathbf{U}_m^{[2]},\mathbf{W}_m^{[2]}\odot\boldsymbol{\lambda}\odot\mathbf{V})|]+\mathrm{Cov}[(\mathbf{U}_m^{[2]},\mathbf{W}_m^{[3]}\odot\boldsymbol{\lambda}\odot\mathbf{V})|]\\ +\mathrm{Cov}[(\mathbf{U}_m^{[3]},\mathbf{W}_m^{[1]}\odot\boldsymbol{\lambda}\odot\mathbf{V})|]&+\mathrm{Cov}[(\mathbf{U}_m^{[3]},\mathbf{W}_m^{[2]}\odot\boldsymbol{\lambda}\odot\mathbf{V})|]+\mathrm{Cov}[(\mathbf{U}_m^{[3]},\mathbf{W}_m^{[3]}\odot\boldsymbol{\lambda}\odot\mathbf{V})|]\}.\end{aligned}$$

Now, let $\bar{\boldsymbol{b}}_{n,m,i}$ be the i-th column of the matrix $\overline{\mathcal{B}}_{n,m}$. Using similar arguments used in YC, we then have

$$\begin{aligned}&\mathrm{Cov}[(\mathbf{U}_n^{[1]},\mathbf{W}_n^{[1]}\odot\boldsymbol{\lambda}\odot\mathbf{V})|]\\ &=\sum_{1\le i\ne j\le r}\sum_{m=1}^{n}(\mathbf{M}^{(i)})'(\overline{\mathbf{B}}_{n,m}^{(i)})'E[(\Delta\mathbf{m}_m^{(i)})'\Delta\mathcal{M}_m|\mathcal{F}_{m-1}]\overline{\mathbf{b}}_{n,m,j}\lambda_j\mathbf{v}_j\\ &=\sum_{1\le i\ne j\le r}\sum_{m=1}^{n}(\mathbf{M}^{(i)})'(\overline{\mathbf{B}}_{n,m}^{(i)})'(u_i\mathbf{v}_i'\mathbf{e}_i-u_i\mathbf{v}_i'\mathbf{u})\overline{\mathbf{b}}_{n,m,j}\lambda_j\mathbf{v}_j\\ &=n\sum_{1\le i\ne j\le r}u_i(\mathbf{M}^{(i)})'\int_0^1\left(\frac{1}{x}\right)^{\overline{(\mathbf{M}^{(i)})'}}\Sigma^{(i)}\left(\frac{1}{x}\right)^{\overline{\mathbf{q}}_j}dx\lambda_j\mathbf{v}_j+o(n),\ \text{a.s.}:=\sum_{1\le i\ne j\le r}u_i\lambda_j(\mathbf{M}^{(i)})'\rho_{i,j}\mathbf{v}_j+o(n),\ \text{a.s.}\end{aligned}$$

$$\begin{aligned}\mathrm{Cov}[(\mathbf{U}_n^{[3]},\mathbf{W}_n^{[3]}\odot\boldsymbol{\lambda}\odot\mathbf{V})|]&=\sum_{1\le i\ne j\le r}\sum_{m=1}^{n}(\mathbf{B}_{n,m}^{(i),1})'(\mathbf{T}^{(i)})'E[(\Delta\mathbf{Q}_m^{(i)})'\Delta\mathcal{Q}_m|\mathcal{F}_{m-1}]\mathbf{H}\mathcal{B}_{n,m}^{[3]}(\mathbf{I}-\mathbf{1}'\mathbf{u})_j\lambda_j\mathbf{v}_j\\ &=\sum_{1\le i\ne j\le r}\sum_{m=1}^{n}(\mathbf{B}_{n,m}^{(i),1})'(\mathbf{T}^{(i)})'(u_i(\sigma_1^2v_{i1},\ldots,\sigma_k^2v_{ik})'\mathbf{e}_i\mathbf{H}\mathcal{B}_{n,m}^{[3]}(\mathbf{I}-\mathbf{1}'\mathbf{u})_j\lambda_j\mathbf{v}_j\\ &=n\sum_{1\le i\ne j\le r}u_i\int_0^1dx\left[\int_x^1\frac{1}{y}\left(\frac{1}{y}\right)^{\overline{\mathbf{M}^{(s)}}}dy\right]'(\mathbf{T}^{(i)})'\Sigma_4^{(i)}\mathbf{H}\left[\int_x^1dy\int_x^ydu\frac{1}{yu}\left(\frac{y}{u}\right)^{\overline{\mathbf{Q}}}\right](\mathbf{I}-\mathbf{1}'\mathbf{u})_j\lambda_j\mathbf{v}_j\\ &:=n\sum_{1\le i\ne j\le r}u_i\lambda_i\boldsymbol{\Lambda}_{i,4}^{\diamond}(\mathbf{I}-\mathbf{1}'\mathbf{u})_j\mathbf{v}_j+o(n)\ \text{a.s.}\end{aligned}$$

$$\begin{aligned}&\mathrm{Cov}[(\mathbf{U}_n^{[1]},\mathbf{W}_n^{[2]}\odot\boldsymbol{\lambda}\odot\mathbf{V})|]=o(n)\ \text{a.s.},\ \mathrm{Cov}[(\mathbf{U}_n^{[1]},\mathbf{W}_n^{[3]}\boldsymbol{\lambda}\mathbf{V})|]=\mathbf{0}\ \text{a.s.},\\ &\mathrm{Cov}[(\mathbf{U}_n^{[2]},\mathbf{W}_n^{[1]}\boldsymbol{\lambda}\mathbf{V})|]=o(n)\ \text{a.s.},\ \mathrm{Cov}[(\mathbf{U}_n^{[2]},\mathbf{W}_n^{[2]}\odot\boldsymbol{\lambda}\odot\mathbf{V})|]=o(n)\ \text{a.s.},\\ &\mathrm{Cov}[(\mathbf{U}_n^{[2]},\mathbf{W}_n^{[3]}\odot\boldsymbol{\lambda}\odot\mathbf{V})|]=\mathbf{0}\ \text{a.s.},\ \mathrm{Cov}[(\mathbf{U}_n^{[3]},\mathbf{W}_n^{[1]}\odot\boldsymbol{\lambda}\odot\mathbf{V})|]=\mathbf{0}\ \text{a.s.},\\ &\mathrm{Cov}[(\mathbf{U}_n^{[3]},\mathbf{W}_n^{[2]}\odot\boldsymbol{\lambda}\odot\mathbf{V})|]=\mathbf{0}\ \text{a.s.}.\end{aligned}$$

Now collecting terms and letting $n\to\infty$, we get

$$\Gamma_{11}=\sum_{1\le i\ne j\le r} u_i\lambda_j[(\mathbf{M}^{(i)})^{'}\rho_{i,j}+\boldsymbol{\Lambda}_{i,4}^{\diamond}(\mathbf{I}-\mathbf{1}^{'}\mathbf{u})_j]\mathbf{v}_j.$$

By similar evaluations we have

$$\Gamma_{14}=\sum_{1\le i\ne j\le r} u_i[(\mathbf{M}^{(i)})^{'}\rho_{i,j}+\boldsymbol{\Lambda}_{i,4}^{\diamond}(\mathbf{I}-\mathbf{1}^{'}\mathbf{u})_j]\mathbf{v}_j,$$
$$\Gamma_{22}=\sum_{1\le i\ne j\le r} u_i[\rho_{i,j}+(\mathbf{I}-\mathbf{1}^{'}\mathbf{v}_i)^{'}\boldsymbol{\Lambda}_{i,4}^{\sharp}(\mathbf{I}-\mathbf{1}^{'}\mathbf{u})_j]\mathbf{v}_j,$$
$$\Gamma_{23}=\sum_{1\le i\ne j\le r} \lambda_i u_i\mathbf{v}_i^{'}[(\rho_{i,j})^{'}+((\mathbf{I}-\mathbf{1}^{'}\mathbf{u})_i)^{'}\boldsymbol{\Lambda}_{j,4}^{\sharp})^{'}(\mathbf{I}-\mathbf{1}^{'}\mathbf{v}_j)].$$

The proof is finished by collecting terms.

Proof of The Corollary

Notice here for fixed s, (s=1,…,r), we assume that $\mathbf{M}^{(s)}=\mathbf{1}_k^{'}\mathbf{v}_s$, $\mathbf{Q}=\mathbf{1}_r^{'}\mathbf{u}$ and that $\boldsymbol{\zeta}_s$ and $\boldsymbol{\theta}$ are constant matrices. Then we have $\overline{\mathbf{M}^{(s)}}=\mathbf{0}$, $\bar{\mathbf{Q}}=\mathbf{0}$, $\sum_1^{(s)}=\{\mathbf{v}_s\}-\mathbf{v}_s^{'}\mathbf{v}_s$, and $(\mathbf{M}^{(s)})^{'}\sum_1^{(s)}\mathbf{M}^{(s)}=\mathbf{0}$. Also, as in the proof of Corollary 3.2 in ZHC and noting that $E(\boldsymbol{\zeta}_s)=\mathbf{M}^{(s)}$ and $E(\boldsymbol{\theta})=\mathbf{Q}$, we have $\boldsymbol{\Lambda}_{s,2}^{\dagger}=\sum_2^{(s)}=E(\boldsymbol{\zeta}_s^{'}\{\mathbf{v}_s\}\boldsymbol{\zeta}_s)-\mathbf{v}_s^{'}\mathbf{v}_s=\mathbf{0}$, $\boldsymbol{\Lambda}_2^{\dagger}=\mathbf{0}$, $\boldsymbol{\Lambda}_{s,3}^{\dagger}=2(\mathbf{T}^{(s)})^{'}\sum_3^{(s)}\mathbf{T}^{(s)}=2\boldsymbol{\Psi}_s$, $\boldsymbol{\Lambda}_3^{\dagger}=2\boldsymbol{\Phi}$, and $\boldsymbol{\Lambda}_{s,23}^{\dagger}=\mathbf{0}$. Similarly,

$$\begin{aligned}
\boldsymbol{\Lambda}_{s,2}^{\sharp}=2\sum\nolimits_2^{(s)},\quad & \boldsymbol{\Lambda}_2^{\sharp}=2\sum\nolimits_2,\ \boldsymbol{\Lambda}_{s,3}^{\sharp}=6(\mathbf{T}^{(s)})^{'}\sum\nolimits_3^{(s)}\mathbf{T}^{(s)}\\
&=6\boldsymbol{\Psi}_s,\ \boldsymbol{\Lambda}_3^{\sharp}=6\boldsymbol{\Phi},\ \boldsymbol{\Lambda}_{s,23}^{\sharp}=\mathbf{0},\ \boldsymbol{\Lambda}_{s,2}^{\diamond}=\mathbf{0},\ \boldsymbol{\Lambda}_2^{\diamond}=\mathbf{0},\ \boldsymbol{\Lambda}_{s,3}^{\diamond}=(\mathbf{T}^{(s)})^{'}\sum\nolimits_3^{(s)}\mathbf{T}^{(s)}\int_0^1\Big(\int_x^1\tfrac{1}{y}dy\Big)\Big(\int_x^1\int_x^y\tfrac{1}{yu}du\,dy\Big)dx\\
&=3(\mathbf{T}^{(s)})^{'}\sum\nolimits_3^{(s)}\mathbf{T}^{(s)}\\
&=3\boldsymbol{\Psi}_s,\ \boldsymbol{\Lambda}_3^{\diamond}=3\boldsymbol{\Phi},\ \boldsymbol{\Lambda}_{s,23}^{\diamond}\\
&=\mathbf{0},\ \boldsymbol{\Lambda}_{s,32}^{\diamond}=\mathbf{0},\ \boldsymbol{\Lambda}_{s,4}^{\sharp}=6\,\mathbf{1}_{k\times k}\boldsymbol{\Gamma}_s\mathbf{1}_{r\times r},\quad \boldsymbol{\Lambda}_{s,4}^{\diamond}\\
&=3\,\mathbf{1}_{k\times k}\boldsymbol{\Gamma}_s\mathbf{1}_{r\times r}
\end{aligned}$$

, and $\rho_{i,j}=\mathbf{1}_{k\times k}\sum^{(i)}\mathbf{1}_r^{'}$.