On Asymptotic Properties for Adaptive Designs in Clinical Trials

Redouane Yagouti
Université Paris-Saclay, Institut Polytechnique de Paris, Inria Lille (Scool Team)
Master in Optimization

September 2025

Supervisors: Rémy Degenne, Emilie Kaufmann
Univ. Lille, Inria, CNRS, Centrale Lille, UMR 9189–CRIStAL, F-59000 Lille, France

Abstract

Adaptive randomization methods in clinical trials aim to balance ethical considerations with statistical efficiency. Classical procedures ensure balance but often suffer from high variability. The Doubly Adaptive Biased Coin Design (DBCD) applies to multiple treatments but does not achieve minimal variance. The Efficient Randomized Adaptive Design (ERADE) improves efficiency through a discontinuous allocation rule and provides asymptotic normality of allocations, but only in the two-treatment case. We propose a new algorithm that generalizes **ERADE** to K treatments, while preserving its key asymptotic properties: almost sure convergence, asymptotic normality, and efficiency. Moreover, the design can be integrated into Track-and-Stop algorithms for Best Arm Identification in multi-armed bandits, providing asymptotic results that complement existing optimality analyses and reinforcing the connection between adaptive clinical trial designs and sequential learning methods.

Contents

Index of Notation

DBCD	Double Adaptive Biased Coin Design
ERADE	Efficient Randomized Adaptive Design
RAR	Response-Adaptive Randomization
MAB	Multi-Armed Bandits
BAI	Best Arm Identification
LIL	Law of the Iterated Logarithm
SLLN	Strong Law of Large Numbers
K	Number of Treatment
n	Number of patients
$X_{m,k}$	Assignment variable, 1 if patient m is assigned treatment k, and 0 otherwise
$\xi_{m,k}$	Response of patient m to treatment k
$N_{m,k}$	Number of patients assigned to treatment k at time m
$[k]$	Set of integers $\{1, \dots, k\}$
Δ_K	The K-dimensional simplex $\{(v_1, \dots, v_K) : v_k \geq 0, \sum_{k=1}^{K} v_k = 1\}$
$\nabla_u(h)$	For a function $h(u, w) : \mathbb{R}^L \times \mathbb{R}^M \to \mathbb{R}^K$, $\nabla_u(h) = \left(\frac{\partial h_k}{\partial u_i}\right)_{i \in [L], k \in [K]}$
$a.s.$	Almost Surely
$i.o.$	Infinitely Often
$A \succeq B$	$A - B$ is positive semi-definite

1 Introduction

1.1 Motivation: clinical trials and Response-Adaptive Randomization (RAR)

Clinical trials are systematic experiments conducted to compare medical treatments under controlled conditions. Patients are enrolled sequentially, assigned to treatments, and their responses are observed. The central challenge is to design allocation rules that avoid bias, preserve unpredictability, and allow for reliable statistical inference. In the earliest trials, patients were often assigned deterministically, for example by strict alternation. While simple, such methods allowed bias. To obtain reliable and unbiased evidence, randomization became the standard method of assigning participants to different treatments.

Over time, different randomization schemes were proposed to address practical and ethical issues. Some aimed at improving balance between groups, such as Efron's (1971) biased coin design, which adjusts the probability of assigning the under-represented treatment to restore balance.

Another class of designs introduced **Response-Adaptive Randomization (RAR)**. Formally, let $(T_1, X_1), \dots, (T_{t-1}, X_{t-1})$ denote the treatment assignments and outcomes of the first $t-1$ patients, where $T_i \in \{A, B\}$ is the treatment indicator and $X_i \in \{0, 1\}$ the binary response (1=success, 0=failure). At time t, the next patient is randomized according to

$$\mathbb{P}(T_t = A \mid \mathcal{F}_{t-1}) = \pi_t, \qquad \mathbb{P}(T_t = B \mid \mathcal{F}_{t-1}) = 1 - \pi_t,$$

where $\mathcal{F}_{t-1} = \{(T_i, X_i) : i = 1, \dots, t-1\}$ is the history of the trial. The probability π_t is updated sequentially to depend on the past responses, typically increasing in favor of the treatment that has shown superior performance so far.

These adaptive schemes are motivated by two complementary goals:

- *ethical*, in that fewer patients are exposed to inferior treatments (reducing regret)
- *statistical*, in that they accelerate the process of identifying the best treatment by concentrating randomization on the superior arm.

Classical examples include Zelen's (1969) Play-the-Winner rule and the urn model of Wei and Durham (1978).

1.2 Randomization and allocation targets

Later theoretical work showed that the variability of treatment allocation is not a technical detail but a central factor in trial performance. For a given target allocation, the statistical power of the final test decreases as allocation variability increases. In other words, two designs may converge to the same target but differ significantly in efficiency if one is more variable.

Hu and Rosenberger (2003) formalized this by proving that efficiency decreases with variability. Hu et al. (2006) further established that variability cannot be reduced arbitrarily as there exists a theoretical lower bound, and designs attaining it are called *asymptotically efficient*.

This perspective reshaped the evaluation of adaptive randomization procedures. It is no longer sufficient for a design to adapt or to satisfy ethical considerations, its asymptotic variability also determines its value.

The **Doubly Adaptive Biased Coin Design (DBCD)**, introduced by Eisele (1994) and Eisele and Woodroofe (1995) and analysed under more general conditions by Hu and Zhang (2004), extends biased coin randomization by adapting both to current imbalances and to sequential parameter estimates.

Let $\hat{\rho}_m = \rho(\hat{\Theta}_m)$ denote the estimated target allocation after m patients, and let $\frac{N_{m,k}}{m}$ be the observed proportion of patients assigned to treatment k up to time m. Then the probability of assigning the $(m+1)$-th patient to treatment k is given by

$$p_{m+1,k} = g\left(\frac{N_{m,k}}{m}, \hat{\rho}_{m,k}\right),$$

where $g(\cdot,\cdot)$ is a function chosen to regulate convergence.

This construction incorporates two levels of adaptation:

- *Adaptation to imbalance:* the dependence on $\frac{N_{m,k}}{m}$ corrects discrepancies between current allocation proportions and the estimated target.
- *Adaptation to parameter estimates:* the target $\hat{\rho}_{m,k}$ itself is updated sequentially as $\hat{\Theta}_m$ evolves with accumulating data.

DBCD ensures almost sure convergence to targets, admits central limit theorems for allocations and estimators, and extends naturally to multiple treatments. However, because it relies on smooth and differentiable allocation functions, its asymptotic variance is larger than the theoretical minimum. Thus, while reliable, DBCD is not asymptotically efficient.

The **Efficient Randomized Adaptive Design (ERADE)**, proposed by Hu et al. (2009), was developed to address this gap. Its innovation is the use of a discontinuous allocation rule. If a treatment is below its target, its assignment probability increases; if above, it decreases; if exactly on target, the assignment follows the target. This step-function adjustment departs from classical smooth functions and is analyzed with martingale stopping-time techniques instead of Taylor expansions. **ERADE** established two key results:

- allocation proportions converge almost surely to the target, with a central limit theorem at rate $n^{-1/2}$,
- the asymptotic variance achieves the *Cramér–Rao* lower bound, meaning **ERADE** minimizes allocation variability.

Thus, **ERADE** is both convergent and efficient. Its limitation, however, is scope: the results are proven only for two treatments, whereas DBCD applies to general K-arm trials. Extending **ERADE** to multiple treatments remains an important open problem, especially since most real trials involve more than two options.

1.3 Multi-armed bandits and Best Arm Identification

The **multi-armed bandit (MAB)** problem is a classical model of sequential decision making under uncertainty, it was introduced as an idealized model for sequential clinical trials. Imagine several experimental drugs, each with an unknown probability of success. Patients arrive sequentially, and for each patient the investigator must choose a treatment. The outcome is observed only for the chosen treatment, while the others remain unknown.

Two objectives are distinguished in the **MAB** literature:

- *Reward maximization:* maximize cumulative outcomes, with performance measured by regret compared to always choosing the best arm.

- *Best Arm Identification (BAI)*: identify, with high confidence, the treatment with the highest mean outcome, while minimizing the number of samples.

For clinical trials, the **BAI** perspective is the natural analogue: the purpose is not to maximize immediate benefit during the trial, but to reliably determine which treatment should be recommended afterward.

The work of Garivier and Kaufmann (2016) shows that asymptotic optimality in **BAI** requires allocating samples according to optimal proportions $\omega^*(\mu)$, defined through an optimization problem involving Kullback–Leibler divergences. The **Track-and-Stop** algorithm achieves this by enforcing discontinuous tracking rules that push empirical allocations back toward these proportions, combined with a stopping rule ensuring correct identification.

This principle is very similar to **ERADE**: both rely on discontinuous adjustments to keep allocations close to their theoretical targets. The main difference is scope: Track in **MAB** context applies naturally to K arms, while **ERADE** is proven only for two.

1.4 Contributions and outline

This work develops a new algorithm that generalizes **ERADE** to the case of K treatments. The procedure extends its discontinuous allocation mechanism beyond two arms while preserving the key asymptotic properties:

- almost sure convergence to target allocations,
- asymptotic normality at the $n^{-1/2}$ rate,
- and efficiency at the variance lower bound.

The proposed algorithm can be viewed as a general class of RAR, that can contain a close version of the **Track-and-Stop** framework for Best Arm Identification proposed by Garivier and Kaufmann (2016). This provides asymptotic normality results for bandit algorithms, offering not only guarantees on sample complexity but also a distributional description of how allocations fluctuate around theoretical targets. Such insights can be seen as complementary, trying to dig into the asymptotic behaviour of **MAB** algorithms, where in bandits literature we focus more on non-asymptotic results.

In this way, the contribution addresses two domains simultaneously: in clinical trials, it provides an efficient multi-treatment response-adaptive design; in **MAB** theory, it enriches the analysis of Track-and-Stop algorithms by adding asymptotic distributional results.

The results presented here were obtained before the appearance of a recent work **?** published in August 2025 (one of the authors of the first 2009 version is also present in the new paper), which extended **ERADE** to the case of multiple treatments. That paper tackles the same open problem raised in Hu et al. (2009), and its timing confirms the growing interest in this direction. Nevertheless, the procedure we propose is more general: it accommodates an arbitrary number of treatments, preserves the asymptotic properties of **ERADE**, and can be embedded into the analysis framework of multi-armed bandits. In particular, this opens the possibility of studying not only convergence and efficiency, but also notions central to the learning literature such as sample complexity and stopping rules.

2 Definitions and Notations

2.1 Common Notations and Setup

We consider a sequential clinical trial with K treatments. Patients arrive sequentially and are assigned to treatments adaptively. To describe the allocation process, we introduce the following notations.

Assignment Vector. For the m-th patient, let

$$X_m = (X_{m,1}, \ldots, X_{m,K})$$

denote the assignment vector, where

$$\forall k \in [K] : \; X_{m,k} = \begin{cases} 1, & \text{if patient } m \text{ receives treatment } k, \\ 0, & \text{otherwise.} \end{cases}$$

This representation encodes the treatment choice as a one-hot vector. It is convenient because cumulative sums of these vectors directly give the number of patients assigned to each treatment, and the same notation applies naturally for general K.

Patient Responses. Let $\xi_{m,k}$ denote the potential response of patient m if assigned to treatment k. In practice, only one of the $\{\xi_{m,k}\}_{k=1}^{K}$ is observed, namely the response corresponding to the treatment actually received.

Remark 2.1. *A more general setup allows $\xi_{m,k}$ to be a d-dimensional vector, $\xi_{m,k} = (\xi_{m,k,1}, \ldots, \xi_{m,k,d}) \in \mathbb{R}^d$, but in this document we restrict attention to the case $d = 1$.*

Cumulative assignments. The total number of patients assigned to each treatment up to time n is

$$N_n = (N_{n,1}, \ldots, N_{n,K}), \qquad N_{n,k} = \sum_{j=1}^{n} X_{j,k}.$$

The normalized vector N_n/n represents the observed allocation proportions at time n. Controlling the long-term behavior of these proportions is the central objective of adaptive designs.

Parameters of interest. For treatment k, the parameter of interest is defined as

$$\theta_k := \mathbb{E}[\xi_{1,k}],$$

which may represent, for instance, a success probability for binary outcomes or a mean response for continuous outcomes. Collectively we write

$$\Theta = (\theta_1, \ldots, \theta_K) \tag{2.1}$$

Target allocation proportions. Let $\mathbb{H} \subset \mathbb{R}$ be an open set and consider a continuous function. For $z \in \mathbb{H}^K$, we denote

$$\rho(z) = (\rho_1(z), \ldots, \rho_K(z)) : \mathbb{H}^K \to [0,1]^K$$

where

$$\forall z \in \mathbb{H}^K : \sum_{k=1}^{K} \rho_k(z) = 1$$

The target allocation proportions is denoted in the following as v defined for the unknown parameters Θ as in (2.1)

$$v = \rho(\Theta) \tag{2.2}$$

Different choices of ρ correspond to different design objectives, such as equal allocation, Neyman allocation for power, or ethically motivated rules.

Sequential estimation. Since Θ is unknown, parameters θ_k are estimated adaptively. With initial guesses $\theta_{0,k}$, the sequential estimator for treatment k at stage m is

$$\hat{\theta}_{m,k} = \frac{\sum_{j=1}^{m} X_{j,k}\xi_{j,k} + \theta_{0,k}}{N_{m,k} + 1}, \qquad \hat{\Theta}_m = (\hat{\theta}_{m,1}, \dots, \hat{\theta}_{m,K}) \tag{2.3}$$

The role of the initial values $\theta_{0,k}$ is to regularize the estimator when few or no patients have yet been assigned to treatment k.

In some situations we may deal with the Bahadur-type representation, which is a more generalized estimator

$$\hat{\theta}_{n,k} = N_{n,k}^{-1} \sum_{j=1}^{n} X_{j,k}\xi_{j,k} + o(N_{n,k}^{-1/2}), \quad \text{as } n \to \infty, \tag{2.4}$$

Plug-in target allocation. At time m, the current estimate of the desired allocation proportions is obtained by substituting $\hat{\Theta}_m$ into ρ:

$$\hat{\rho}_m = \rho(\hat{\Theta}_m).$$

Adaptive randomization procedures compare the current allocation proportions N_m/m with this estimated target $\hat{\rho}_m$, and adjust the probability of assigning the next patient accordingly.

In the following we will use the notations $o_p(\cdot)$ and $O_p(\cdot)$ are probabilistic analogues of the classical little-o and big-O notations in analysis.

- $X_n = o_p(a_n)$ means that X_n is *negligible compared to a_n in probability*, i.e. the ratio X_n/a_n converges to 0 in probability.
- $X_n = O_p(a_n)$ means that X_n is *bounded in probability relative to a_n*, i.e. the ratio X_n/a_n does not diverge and remains stochastically bounded.

Formal definitions are given below.

Definition 2.1. *Let $\{X_n\}$ be a sequence of random variables and $\{a_n\}$ a sequence of positive real numbers.*

- *We say that $X_n = o_p(a_n)$ if for every $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{|X_n|}{a_n} > \varepsilon\right) = 0.$$

- *We say that $X_n = O_p(a_n)$ if for every $\varepsilon > 0$ there exists $M > 0$ such that*

$$\limsup_{n\to\infty} \mathbb{P}\left(\frac{|X_n|}{a_n} > M\right) \le \varepsilon.$$

2.2 Examples of Target Functions ρ

The function ρ determines the desired allocation proportion as a function of the unknown parameters of the treatment responses. In the literature, several choices have been proposed depending on the nature of the responses and the optimality criteria. We illustrate below some commonly used examples from Hu et al. (2009).

Example 2.1 (Efron's Biased Coin Design)**.** *For balancing two treatments with equal target allocation $\rho = 1/2$, **ERADE** assigns the $(m+1)$th patient to treatment 1 with probability*

$$p_{m+1} = \begin{cases} \alpha/2, & \text{if } N_{m,1}/m > 1/2, \\ 1/2, & \text{if } N_{m,1}/m = 1/2, \\ 1-\alpha/2, & \text{if } N_{m,1}/m < 1/2, \end{cases}$$

where $0 \le \alpha < 1$ controls the degree of randomization. The special case $\alpha = 2/3$ corresponds to Efron's classical biased coin design.

Example 2.2 (Binary Response)**.** *Suppose treatment k has success probability P_k and failure probability $Q_k = 1 - P_k$. Natural estimates are*

$$\hat{P}_{m,k} = \frac{\sum_{j=1}^{m} X_{j,k}\xi_{j,k} + 0.5}{N_{m,k}+1}, \quad k = 1, 2,$$

where $\xi_{j,k}$ is the response indicator. Based on these estimates, several target functions $\rho(P_1, P_2)$ have been proposed:

$$v_1 = \frac{1-P_1}{2-P_1-P_2}, \qquad \text{Urn allocation proportion,} \tag{2.5}$$

$$v_2 = \frac{\sqrt{P_1}}{\sqrt{P_1}+\sqrt{P_2}}, \qquad \text{Optimal allocation minimizing failures,} \tag{2.6}$$

$$v_3 = \frac{\sqrt{P_1 Q_1}}{\sqrt{P_1 Q_1}+\sqrt{P_2 Q_2}}, \qquad \text{Neyman allocation.} \tag{2.7}$$

These examples illustrate how the **ERADE** framework accommodates a wide range of allocation targets, from balance-driven rules to efficiency-based optimal criteria.

2.3 Why Convergence to Target Matters?

The goal of adaptive randomization is to ensure that the observed allocation proportions

$$\frac{N_n}{n} = \left(\frac{N_{n,1}}{n}, \ldots, \frac{N_{n,K}}{n}\right)$$

converge almost surely to the target allocation $\rho(\Theta)$. The vector $\rho(\Theta)$ encodes the design objective, which may correspond to balance between treatments, optimal power for statistical tests, or ethical considerations such as assigning more patients to superior treatments. If the convergence

$$\frac{N_n}{n} \longrightarrow \rho(\Theta) \quad \text{a.s.}$$

fails to hold, then the actual allocation may deviate from the intended target, undermining both the ethical and statistical validity of the trial. Establishing almost sure convergence of $N_n/n \to \rho(\Theta)$ is therefore the fundamental property of any adaptive design, and provides the basis for further asymptotic analysis such as central limit theorems and efficiency comparisons.

3 Doubly Biased Coin Designs (DBCD), Hu and Zhang (2004)

In a multi-treatment clinical trial with $K \geq 2$ treatments, the **Doubly Adaptive Biased Coin Design (DBCD)** is defined as follows. Patients are enrolled sequentially, and at each step $m+1$, the assignment depends both on the current allocation proportions and on sequential estimates of the desired target allocation.

3.1 Allocation rule

The DBCD specifies an *allocation function*

$$g(x,y) = (g_1(x,y), \ldots, g_K(x,y)) : [0,1]^K \times [0,1]^K \to [0,1]^K,$$

with $\sum_{k=1}^K g_k(x,y) = 1$. At stage $m+1$, the next patient is assigned to treatment k with probability

$$p_{m+1,k} = g_k\left(\frac{N_m}{m}, \hat{\rho}_m\right), \quad k = 1, \ldots, K.$$

Thus the design adapts in two ways:

- to the current imbalance between observed proportions N_m/m and the target $\hat{v}_m$,
- to updated estimates of the target $\hat{v}_m$ itself as more responses are observed.

Algorithm 1 Adaptive Allocation Algorithm DBCD

$\mathbb{P}(X_{1,j} = 1) = \frac{1}{K}, \quad 1 \leq j \leq K$
$\hat{\Theta}_1 = (\theta_{0,j})_{1 \leq j \leq K}$
$N_1 = X_1$
for $m = 2, \ldots, n$ **do**
  $\hat{\rho}_{m-1} = \rho(\hat{\Theta}_{m-1})$
  $\mathbb{P}(X_{m,j} = 1 \mid \mathcal{F}_{m-1}) = g\left(\frac{N_{m-1}}{m-1}, \hat{\rho}_{m-1}\right)$
  $N_m = N_{m-1} + X_m$
  $\xi_{m,j} \sim \nu_j(\theta_j), \quad 1 \leq j \leq K$
  $\hat{\theta}_{m,j} = \frac{\sum_{i=1}^m X_{i,j}\xi_{i,j} + \theta_{0,j}}{N_{m,j}+1}$
end for

Example 3.1 (Allocation functions)**.** *A standard family of allocation functions proposed in Hu and Zhang (2004) is:*

$$g_k^{(\gamma)}(x,y) = \frac{y_k \left(\frac{y_k}{x_k}\right)^\gamma}{\sum_{j=1}^K y_j \left(\frac{y_j}{x_j}\right)^\gamma}, \quad \gamma \geq 0, \quad k = 1, \ldots, K, \tag{3.1}$$

with the conventions $g_k^{(\gamma)}(0,y) > 0$ when $x_k = 0$.

Interpretation.

- When $\gamma = 0$, $g_k^{(0)}(x,y) = y_k$, yielding the *adaptive randomized design* of Melfi, Page and Geraldes (2001).

- As $\gamma \to \infty$, the design becomes nearly deterministic, forcing allocations to follow the target vector y closely.
- For intermediate γ, the design balances randomness with variability: larger γ reduces allocation variance while still preserving randomization.

This family generalizes Efron's biased coin idea and ensures convergence of allocations to the desired proportions.

3.2 Conditions for Asymptotic Analysis

Conditions A. The response sequence $\{\xi_n = (\xi_{n,1}, \dots, \xi_{n,K})\}$ satisfies the following conditions

(A1) $\forall k \in [K]: \quad \mathbb{E}|\xi_{1,k}| < \infty$.

(A2) There exists $\varepsilon > 0$ such that $\forall k \in [K]: \quad \mathbb{E}|\xi_{1,k}|^{2+\varepsilon} < \infty$.

Conditions B. The allocation function $g(x, y)$ satisfies the following

(B1) $g(v, v) = v$, and $g(x, y) - g(x, v) \to 0$ uniformly in x as $y \to v$ for $x.\mathbf{1}' = y.\mathbf{1}' = 1$

(B2) There exists $0 \le \lambda_0 < 1$ such that for each $k \in [K]$

$$\frac{g_k(x, v) - g_k(v, v)}{x_k - v_k} \le \lambda_0, \quad \text{for } x_k > v_k \text{ and } x.\mathbf{1}' = 1$$

(B3) For any $0 < \epsilon < 1/K$, there exists $c_\epsilon > 0$ satisfying

$$g_k(x, y)\,|_{x_k=0} \ge c_\epsilon \quad \text{with } x.\mathbf{1}' = y.\mathbf{1}' = 1,\ y \in [\epsilon, 1)^K \tag{3.2}$$

$$\liminf_{x_k \to 0^+} \frac{g_k(x, y)}{\min\{x_1, \dots, x_K\}} \ge c_\epsilon \quad \text{with } x.\mathbf{1}' = y.\mathbf{1}' = 1,\ y \in [\epsilon, 1)^K \tag{3.3}$$

(B4) $g(x, y)$ admits local linear approximation at (v, v), there exists $\delta \in (0, 1)$ for which

$$\begin{aligned} g(x, y) = g(v, v) &+ \sum_{k=1}^{K} (x_k - v_k) \frac{\partial g}{\partial x_k}\bigg|_{(v,v)} + \sum_{k=1}^{K} (y_k - v_k) \frac{\partial g}{\partial y_k}\bigg|_{(v,v)} \\ &+ o(\|x - v\|^{1+\delta} + \|y - v\|^{1+\delta}) \qquad \text{when } (x, y) \to (v, v) \end{aligned}$$

Conditions C. The proportion function $\rho(z)$ satisfies the conditions

(C1) $\rho(z)$ is continuous and $\rho(\Theta) = v$.

(C2) $\rho(z)$ admits a local linear approximation around Θ, there exists $\delta > 0$ such that

$$\rho(z) = \rho(\Theta) + \sum_{k=1}^{K} (z_k - \theta_k) \frac{\partial \rho}{\partial z_k}\bigg|_{\Theta} + o(\|z - \Theta\|^{1+\delta}), \qquad \text{as} \quad z \to \Theta \tag{3.4}$$

Remark 3.1. *If conditions (B4) and (C2) hold, then*

$$g\,(x, \rho(z)) = g(v, v) + (x - v)\,H + (z - \Theta)\,\nabla(\rho)\big|_{\Theta}\,E + o\big(\|x - v\|^{1+\delta}\big) + o\big(\|z - \Theta\|^{1+\delta}\big), \tag{3.5}$$

as $(x, z) \to (v, \Theta)$, where

$$H = \nabla_x g\big|_{(v,v)} \in \mathbb{R}^{K \times K}, \qquad E = \nabla_\rho g\big|_{(v,v)} \in \mathbb{R}^{K \times K}, \tag{3.6}$$

Spectral quantities λ and ν. Since $g(x,y)\mathbf{1}' = 1$, we have $H\mathbf{1}' = E\mathbf{1}' = 0'$. Thus 0 is an eigenvalue of H. Let $T^{-1}HT = \operatorname{diag}\big[0, J_2, \ldots, J_s\big]$ be the Jordan decomposition, where each J_t is a Jordan block of size ν_t with eigenvalue λ_t. Define

$$\lambda \;=\; \max\{\Re(\lambda_2), \ldots, \Re(\lambda_s)\}, \qquad \nu \;=\; \max\{\nu_j : \Re(\lambda_j) = \lambda\}. \tag{3.7}$$

These quantities govern the growth bounds that appear repeatedly later; for instance, $\|B_{n,m}\| \leq C\,(n/m)^{\lambda}\,\log^{\nu-1}(n/m)$ in the iterative expansions based on H.

3.3 Asymptotic Properties

Lemma 3.1 (Lemma A.4 Hu and Zhang (2004))**.** *For each $k = 1, \ldots, K$, if $\{N_{n,k} \to \infty\}$ a.s. then we have*

$$\hat{\theta}_{n,k} \to \theta_k \qquad \text{, if (A1)} \tag{3.8}$$

$$\hat{\theta}_{n,k} - \theta_k = O\left(\sqrt{\frac{\log\log N_{n,k}}{N_{n,k}}}\right) \qquad \text{, if (A2)} \tag{3.9}$$

Proof. In his paper, Doob (1936) explains how subsequences selected following a rule from a sequence of random variables behaves. It states that starting from an infinite sequence of *iid* $X_1, X_2, \ldots$ if we consider a rule of selecting these variables that only depends on the past observations and if this rule picks infinitely many variables almost surely, then the resulting subsequence is iid with identical distribution.

For $k \in [K]$ let $\tau_i^k := \min\{j; N_{j,k} = i\}$ the first time treatment k is picked i times. Let $\{\eta_{m,k}\}$ a copy of $\{\xi_{m,k}\}$ for $m = 1, 2, \ldots$ and consider the following selection rule

$$\Xi_{m,k} = \xi_{\tau_m^k,k}\mathbb{1}_{\{\tau_m^k < +\infty\}} + \eta_{m,k}\mathbb{1}_{\{\tau_m^k = +\infty\}}$$

From Doob (1936) we can show that $\{\Xi_{m,k}\}$ is a sequence of i.i.d random variable with same distribution of $\xi_{1,k}$. Then from LLN and LIL we have

$$n^{-1}\sum_{m=1}^{n}\Xi_{m,k} \to \theta_k \quad \text{and} \quad \sum_{m=1}^{n}\Xi_{m,k} - n\theta_k = O(\sqrt{n\,\mathrm{loglog}\,n}) \quad \text{a.s.}$$

note that from (2.3) we can rewrite our estimator as $\hat{\theta}_{n,k} = \frac{1}{N_{n,k}+1}\left(\sum_{m=1}^{N_{n,k}}\Xi_{m,k} + \theta_{0,k}\right)$, so in the event $\{N_{n,k} \to +\infty\}$ we conclude (3.8) and (3.9). ☐

Lemma 3.2 (Lemma A.5 Hu and Zhang (2004))**.** *For* ***DBCD*** *design, if conditions (A1), (B3) and (C1) are satisfied, then for $k \in [K]$ we have*

$$N_{n,k} \to +\infty \quad a.s. \tag{3.10}$$

and we have

$$\hat{\Theta}_n \to \Theta \quad and \quad \hat{\rho}_n \to v \quad a.s. \tag{3.11}$$

If, in addition, (A2) holds, we further have:

$$\hat{\theta}_{n,k,i} - \theta_{k,i} = O\left(\sqrt{\frac{\log\log N_{n,k}}{N_{n,k}}}\right) \quad a.s. \tag{3.12}$$

Proof. **Step 1:** If $\lim_{n\to+\infty} N_{n,k} < +\infty$, then $\{\hat{\theta}_{n,k}\}$ takes finitely many values. Else, if $\lim\limits_{n\to+\infty} N_{n,k} = +\infty$, from Lemma 3.1, it follows that $\hat{\theta}_{n,k} \to \theta_k$ *a.s.* So in any case: $\hat{\Theta}_n \to \Theta$.

The closure of $\{\hat{\Theta}_n\}$ given by $Cl(\{\Theta_n\}) = \{\hat{\Theta}_n\} \cup \{\text{accumulation points}\}$ is compact (bounded and finite dimension), so $\{\hat{\Theta}_n\}$ is relatively compact. Since $\forall y \in Cl(\{\Theta_n\}) : \rho(y) \in (0,1)^K$, using the continuity of ρ we have

$$\exists\, 0 < \delta \leq \frac{1}{K} : \quad \hat{\rho}_n = \rho(\hat{\Theta}_n) \in [\delta, 1)^K.$$

Step 2: On the event $\{N_{n,j} = 0,\ n = 1, 2, \dots\}$, we can use the first condition of (B3) resulting $\mathbb{P}(X_{n,j} = 1 \mid \mathcal{F}_{n-1}) \geq c_\delta$, then

$$\sum_{n=2}^{+\infty} \mathbb{P}(X_{n,j} = 1 \mid \mathcal{F}_{n-1}) = +\infty.$$

Using the extended Borel–Cantelli Lemma C.1 we obtain that $\{X_{n,j} = 1 \ i.o.\} = \{N_{n,j} \to +\infty\} \quad a.s.$, which is a contradiction. Therefore:

$$\lim_{n \to +\infty} N_{n,j} \geq 1 \quad a.s. \quad j = 1, \dots, K.$$

Step 3: On the event $\{\lim_{n \to +\infty} N_{n,j} < +\infty\}$, we use (B3) second condition resulting for n large enough:

$$\begin{aligned} \mathbb{P}(X_{n,j} = 1 \mid \mathcal{F}_{n-1}) &\geq \frac{c_\delta}{2} \min\left\{\frac{N_{n-1,j}}{n-1},\ j = 1, \dots, k\right\} \\ &\geq \frac{c_\delta}{2} \cdot \frac{1}{n-1} \end{aligned}$$

Hence, we obtain $\sum_{n=2}^{+\infty} \mathbb{P}(X_{n,j} = 1 \mid \mathcal{F}_{n-1}) = +\infty$, again with Lemma C.1 we obtain that $\{X_{n,k} = 1 \ i.o.\} = \{N_{n,k} \to +\infty\}$ a.s. which is a contradiction. Therefore, we conclude:

$$\lim_{n \to +\infty} N_{n,k} = +\infty \quad \text{a.s.} \quad k = 1, \dots, K.$$

So using the continuity of ρ and Lemma 3.1 we can conclude on the proof.

□

Remark 3.2. *The Lemma 3.2 ensures that under mild assumptions, the number of patients assigned to each treatment arm grows without bound almost surely. In other words, each arm continues to receive allocations infinitely often. Condition (B3) plays a central role here as it enforces that whenever the allocation proportion for a treatment is too small relative to its target, the allocation rule must give that treatment a strictly positive probability of being assigned.*

The following result guarantees that the design achieves its intended allocation goal no matter how randomness plays out in the early stages of the trial, the long-run proportion of patients assigned to each treatment converges to the pre-specified target $v = \rho(\Theta)$. This strong convergence property is essential before one can study more refined asymptotic results such as central limit theorems and efficiency bounds.

Theorem 3.1 (Theorem 4.1 Hu and Zhang (2004))**.** *If conditions (A1), (B1), (B2), (B3) and (C1)are satisfied, then:*

$$\hat{\rho}_n \to v \quad a.s. \tag{3.13}$$

$$\frac{N_n}{n} \to v \quad a.s. \tag{3.14}$$

Proof. Under conditions (A1), (B3) and (C1), Lemma 3.2 holds, thus $\hat{\rho}_n \to v$..

Let $M_n = \sum_{m=1}^{n} \Delta M_m \quad$ with $\quad \Delta M_m = X_m - \mathbb{E}[X_m \mid \mathcal{F}_{m-1}]$.

Then we have:

$$\begin{aligned}\Delta(N_n - nv) &= X_n - v\\ &= \Delta M_n + g\left(\frac{N_{n-1}}{n-1}, \hat{\rho}_{n-1}\right) - v\\ &= \Delta M_n + g\left(\frac{N_{n-1}}{n-1}, \hat{\rho}_{n-1}\right) - g\left(\frac{N_{n-1}}{n-1}, v\right) + g\left(\frac{N_{n-1}}{n-1}, v\right) - g(v,v).\end{aligned}$$

Let $k = 1, \dots, K$ and define $q_{n,k}$ and $p_{n,k}$ as follow

$$\begin{aligned}q_{n,k} &= N_{n,k} - nv_k\\ \Delta p_{n,k} &= \Delta M_{n,k} + g_k\left(\frac{N_{n-1}}{n-1}, \hat{\rho}_{n-1}\right) - g_k\left(\frac{N_{n-1}}{n-1}, v\right)\end{aligned}$$

$\{M_{n,k}\}_{n\geq 0}$ being a martingale we have $M_{n,k} = o(n)$ a.s., and condition (B1) gives that $g_k\left(\frac{N_{n-1}}{n-1}, \hat{\rho}_{n-1}\right) - g_k\left(\frac{N_{n-1}}{n-1}, v\right) = o(1)$, so we obtain:

$$p_{n,k} = \sum_{m=1}^{n} p_{m,k} = M_{n,k} + \sum_{m=1}^{n} o(1) = o(n)$$

On the other hand:

$$\begin{aligned}\Delta q_{n,k} &= \Delta(N_{n,k} - nv_k)\\ &= \Delta p_{n,k} + g_k\left(\frac{N_{n-1}}{n-1}, v\right) - g_k(v,v)\\ &= \Delta p_{n,k} + \frac{g_k\left(\frac{N_{n-1}}{n-1}, v\right) - g_k(v,v)}{\frac{N_{n-1,k}}{n-1} - v_k} \cdot \frac{q_{n-1,k}}{n-1}\\ \implies q_{n,k} &= \Delta p_{n,k} + q_{n-1,k} + \underbrace{\frac{g_k\left(\frac{N_{n-1}}{n-1}, v\right) - g_k(v,v)}{\frac{N_{n-1,k}}{n-1} - v_k}}_{\leq \lambda_0;\text{ by condition (B2)}} \cdot \frac{q_{n-1,k}}{n-1}\end{aligned}$$

If $q_{n-1,k} \geq 0$ then $q_{n,k} \leq \left(1 + \frac{\lambda_0}{n-1}\right) q_{n-1,k} + \Delta p_{n,k}$. Otherwise, if $q_{n-1,k} \leq 0$ then there exists $k_1 \geq 0$ such that $q_{n,k} = q_{n-1,k} + X_{n,k} - v_k \leq |X_{n,k} - v_k| \leq k_1$

And for the term $\Delta p_{n,k}$ there exists $k_2 \geq 0$ such that $|\Delta p_{n,k}| \leq k_2$. Thus, for $k_0 = k_1 + k_2$ we obtain

$$q^+_{n,k} \leq \left(1 + \frac{\lambda_0}{n-1}\right)\left(q^+_{n-1,k} \vee k_1\right) + \Delta p_{n,k}$$

Using Lemma A.3 from Hu and Zhang (2004), we have:

$$\begin{aligned}q^+_{n,k} &\leq C\left\{\sum_{m=1}^{n} \frac{|p_{m,k}|}{m}\left(\frac{n}{m}\right)^{\lambda_0} + \max_{1\leq m\leq n}\left(\frac{n}{m}\right)^{\lambda_0}|p_{m,k}| + (k_0 + |q_{1,k}|)n^{\lambda_0}\right\}\\ &= \sum_{m=1}^{n} \frac{O(m)}{m}\left(\frac{n}{m}\right)^{\lambda_0} + \max_{1\leq m\leq n}\left(\frac{n}{m}\right)^{\lambda_0} O(m) + O(n^{\lambda_0}), \quad \lambda_0 \in [0,1)\\ &= o(n)\end{aligned}$$

So we conclude that $\forall k \in [K] : \limsup_{n\to+\infty} \frac{N_{n,k}}{n} - v_k \leq 0$. Now using Bolzano–Weierstrass, let $\bar{v}$ be an adherence value of $\frac{N_n}{n}$, then

$$\begin{aligned}&\bar{v}_k - v_k \leq 0, \quad k = 1, \dots, K, \quad \text{with} \quad \bar{v}\mathbf{1}' = v\mathbf{1}' = 1.\\ &\implies \bar{v} = v.\end{aligned}$$

Thus, we have proven (3.14) $\square$

Theorem 3.2 (Theorem 4.2 Hu and Zhang (2004)). *Suppose* $\frac{N_n}{n} \to v$ *a.s. and conditions (A2) , (B4) and (C2) hold. Suppose* $\lambda < 1$, *then for all* $\kappa > \frac{1}{2} \vee \lambda$:

$$\hat{\rho}_n - v = O\left(\sqrt{\frac{\log\log(n)}{n}}\right) \quad a.s. \tag{3.15}$$

$$n^{-\kappa}(N_n - nv) \to 0 \quad a.s. \tag{3.16}$$

Furthermore, if $\lambda < \frac{1}{2}$, *then:*

$$N_n - nv = O\left(\sqrt{n\log\log(n)}\right) \quad a.s. \tag{3.17}$$

Proof. By assumption $\frac{N_n}{n} \to v$ a.s. with $v \in (0,1)$ so $N_n \to +\infty$ a.s. So by Lemma 3.1 we have that $\hat{\Theta}_n - \Theta = o\left(\sqrt{\frac{\log\log(n)}{n}}\right)$ *a.s.*. Injecting this in condition (C2) we obtain

$$\begin{aligned}\hat{\rho}_n - v &= \rho(\hat{\Theta}_n) - \rho(\Theta) \\ &= O\left(\sqrt{\frac{\log\log(n)}{n}}\right) \quad a.s.\end{aligned}$$

which proves (3.15).

Using condition (B4), we have the expansion:

$$\begin{aligned}\Delta(N_n - nv) &= X_n - v \\ &= \Delta M_n + g\left(\frac{N_{n-1}}{n-1}, \hat{\rho}_{n-1}\right) - g(v,v) \\ &= \Delta M_n + \left(\frac{N_{n-1}}{n-1} - v\right)\cdot H + (\hat{\rho}_{n-1} - v)\cdot E + o\left(\left\|\frac{N_{n-1}}{n-1} - v\right\|^{1+\delta}\right) + o(\|\hat{\rho}_n - v\|^{1+\delta}) \\ &= \left(\frac{N_{n-1}}{n-1} - v\right)\cdot H + o\left(\left\|\frac{N_{n-1}}{n-1} - v\right\|\right) + \Delta M_n + O(\|\hat{\rho}_n - v\|) \quad a.s.\end{aligned} \tag{3.18}$$

By Theorem C.1 applied on $M_n = O\left(\sqrt{n\log\log(n)}\right)$ *a.s.* Thus, using (3.15) we have:

$$\begin{aligned}M_n + \sum_{m=1}^{n} O(\|\hat{\rho}_n - v\|) &= M_n + \sum_{m=1}^{n} O\left(\sqrt{\frac{\log\log(m)}{m}}\right) \\ &= O\left(\sqrt{n\log\log(n)}\right) \quad a.s.\end{aligned} \tag{3.19}$$

So using (3.18), (3.19) and Lemma A.2 we obtain

$$\begin{aligned}\forall \beta > 0: \; N_n - nv &= \sum_{m=1}^{n} O\left(\sqrt{m\log\log(m)}\right)\frac{1}{m}\left(\frac{n}{m}\right)^{\lambda+\beta} \\ &= \begin{cases} O\left(\sqrt{n\log\log(n)}\right), & \text{if } \lambda+\beta < \frac{1}{2}, \\ O\left(n^{\lambda+\beta}\right), & \text{if } \lambda+\beta > \frac{1}{2}. \end{cases}\end{aligned}$$

and by choosing δ arbitrarily small we prove (3.16) and (3.17) □

The following result establishes the asymptotic normality of the DBCD, which goes beyond the previously stated theorems. The key result is that both the empirical allocation proportions $\frac{N_n}{n}$ and the estimated targets $\hat{\rho}_n$ fluctuate at the order $n^{-1/2}$ and converge jointly to a multivariate normal distribution. To state the next result we need to define some quantities

- $V_k = Var\,(\xi_{1,k})$ for $k \in [K]$
- $V = \text{diag}\Big(\frac{1}{v_1}V_1, \dots, \frac{1}{v_K}V_K\Big)$
- $\Sigma_1 = \text{diag}(v) - v^\top v$
- $\Sigma_3 = \big(\nabla(\rho)\big|_\Theta\big)^\top V\, \nabla(\rho)\big|_\Theta,$
- $\Sigma_2 \; = \; E^\top \Sigma_3 E$

Theorem 3.3 (Theorem 4.3 in Hu and Zhang (2004))**.** *Assume $N_n/n \to v$ a.s. and that (A2), (B2), (B4) and (C2) hold and that $\lambda < \frac{1}{2}$. Then*

$$\sqrt{n}\left(\tfrac{N_n}{n} - v\,,\, \hat{\rho}_n - v\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,\; \Sigma)\,, \qquad \Sigma \;=\; \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_3 \end{pmatrix}, \tag{3.20}$$

where

$$\Sigma_{11} \;=\; \frac{\Sigma_1}{1-2\lambda} \;+\; \frac{2\,\Sigma_2}{(1-\lambda)(1-2\lambda)}, \qquad \Sigma_{12}^\top \;=\; \Sigma_{21} \;=\; \frac{1}{1-\lambda}\,\Sigma_3\, E,$$

Proof. See Appendix A □

Remark 3.3. *This asymptotic distribution is the foundation for comparing adaptive designs in terms of efficiency. In particular, one can ask whether the asymptotic variance attains the theoretical lower bound. For DBCD, the question of the variance is larger than the minimum is still to be investigated. This shortcoming is one if the main motivations of the development of* ***ERADE****.*

4 Efficient randomized-adaptive design (ERADE), Hu et al. (2009)

The DBCD framework relies on allocation functions $g(x, y)$ that are continuous in their arguments. Hu et al. (2009) proposed the Efficient Randomized-Adaptive Design (ERADE), which replaces the smooth allocation function of DBCD by a simple discontinuous rule. The key idea is that whenever the current allocation proportion deviates from the target, the next assignment probability is shifted roughly rather than smoothly, enforcing a closer tracking of the target.

4.1 Allocation Rule Definition

Let m_0 subjects be initially assigned to each treatment by restricted randomization. After $m > 2m_0$ patients have been treated and their responses observed, denote by $\hat{\rho}_m$ the estimated target allocation. The **ERADE** assigns the $(m+1)^{\text{th}}$ patient to treatment 1 with probability

$$p_{m+1} = \begin{cases} \alpha\hat{\rho}_m, & \text{if } N_{m,1}/m > \hat{\rho}_m, \\ \hat{\rho}_m, & \text{if } N_{m,1}/m = \hat{\rho}_m, \\ 1-\alpha(1-\hat{\rho}_m), & \text{if } N_{m,1}/m < \hat{\rho}_m, \end{cases} \tag{4.1}$$

where $0 \leq \alpha < 1$.

Equivalently, the rule can be expressed in terms of a DBCD allocation function

$$g(x, y; \alpha) = \alpha y\, \mathbf{1}_{x>y} + y\, \mathbf{1}_{x=y} + \big(1 - \alpha(1-y)\big)\, \mathbf{1}_{x<y}, \tag{4.2}$$

Remark 4.1. *The **ERADE** can be seen as a special case of the DBCD framework, obtained by selecting the discontinuous allocation function g given in* (4.2)*. This represents a fundamental departure from earlier designs such as DBCD, where the allocation function g was assumed to be continuous and differentiable.*

*By contrast, **ERADE** uses a piecewise constant allocation function, which requires far fewer technical assumptions but a whole different analysis than **DBCD**. The rule depends only on whether the current allocation proportion lies above, below, or exactly at the target, and not on smooth variations around it.*

Algorithm 2 Adaptive Allocation **ERADE**, $K = 2$ arms (Hu et al. (2009))

---

for $i \in \mathrm{range}(m_0)$ **do**
 $X_{i,i} = 1$; $X_{m_0+i,i} = 1$; $\xi_{i,1} \sim \nu_i(\theta_1)$; $\xi_{m_0+i,2} \sim \nu_i(\theta_2)$
end for
$N_{2m_0} = (2, 2)$
$\hat{\Theta}_{2m_0,k} = \frac{\sum_{i=1}^{2m_0} X_{i,k}\xi_{i,k}+\theta_{0,k}}{N_{2m_0,k}+1}$
for $m \in \mathrm{range}(2m_0, n)$ **do**
 $\hat{\rho}_m = \rho(\hat{\Theta}_m)$
 $\hat{p}_{m+1} = \alpha\hat{\rho}_m \mathbf{1}\left\{\frac{N_m}{m} > \hat{\rho}_m\right\} + \hat{\rho}_m \mathbf{1}\left\{\frac{N_m}{m} = \hat{\rho}_m\right\} + (1-\alpha(1-\hat{\rho}_m))\,\mathbf{1}\left\{\frac{N_m}{m} < \hat{\rho}_m\right\}$
 $\mathbb{P}(X_{m+1,j} = 1 \mid \mathcal{F}_m) = \hat{p}_{m+1}$
 $N_{m+1} = N_m + X_{m+1}$
 $\xi_{m+1,j} \sim \nu_j(\theta_j), \quad 1 \le j \le K$
 $\hat{\theta}_{m+1,j} = \frac{\sum_{i=1}^{m+1} X_{i,j}\xi_{i,j}+\theta_{0,j}}{N_{m+1,j}+1}$
end for

4.2 Theoretical Results

Before discussing consequences, we first record the basic large-sample behavior of **ERADE** under its discontinuous allocation rule g. Unlike smooth designs in DBCD, the analysis proceeds via martingale arguments and yields both almost sure convergence to the target and a central limit theorem.

Let

$$V = \mathrm{diag}\left(\frac{\sigma_1^2}{v}, \ \frac{\sigma_2^2}{1-v}\right),$$

where $\sigma_k^2 = Var(\xi_{1,k})$ for $k = 1, 2$. Define

$$\sigma^2 = \big(\nabla\rho(\Theta)\big)^\top V\,\nabla\rho(\Theta).$$

Theorem 4.1 (Theorem 2.1 in Hu et al. (2009))**.** *Under conditions **A** and **B** we have*

(i) For every $k \in [K]$ as $n \to \infty$,

$$\begin{aligned}
|N_{n,1} - n\hat{\rho}_n| &= o_P(\sqrt{n}) \quad a.s. \\
|N_{n,1} - n\hat{\rho}_n| &= O(\sqrt{n\log\log n}) \quad a.s. \\
N_{n,1} - nv &= O(\sqrt{n\log\log n}) \quad a.s. \\
\hat{\Theta}_n - \Theta &= O\Big(\sqrt{\frac{\mathrm{loglog}(n)}{n}}\Big) \quad a.s.
\end{aligned}$$

(ii) The asymptotic normality results

$$\begin{aligned}
\sqrt{n}(\hat{\Theta}_n - \Theta) &\xrightarrow{\mathcal{D}} N(0, V) \\
\sqrt{n}\left(\frac{N_{n,1}}{n} - v\right) &\xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2).
\end{aligned}$$

Proof. Generalised proof is given in section 5.2 □

Remark 4.2. *The first part shows* ***ERADE*** *achieves its intended target almost surely. The second part gives a distributional description of the residual fluctuations, the allocation proportion behaves like a mean-zero Gaussian perturbation of order* $n^{-1/2}$.

The following result establishes the asymptotic efficiency of **ERADE** achieving the Cramér–Rao lower bound. In practical terms, this means that no other randomized response-adaptive design can produce smaller first-order variability in the allocation proportions.

Theorem 4.2 (Theorem 2.2 in Hu et al. (2009))**.** *Under conditions* ***A*** *and* ***B****, if* $Var(\xi_{1,k}) = I_k^{-1}$ *for* $k = 1, 2$*, where* I_k *is the Fisher information matrix for* θ_k*, then the asymptotic variance of* $\frac{N_{n,1}}{\sqrt{n}}$ *for the* ***ERADE*** *attains the Cramér–Rao lower bound:*

$$\left(\frac{\partial \rho}{\partial y} \bigg|_{\Theta} \right)^{\top} \mathrm{diag}\big((vI_1)^{-1}, ((1-v)I_2)^{-1}\big) \left(\frac{\partial \rho}{\partial y} \bigg|_{\Theta} \right) \tag{4.3}$$

5 Extended ERADE with K Treatments

5.1 Definition and Setup

Let m_0 subjects be initially assigned to each treatment by restricted randomization. After $m > K\,m_0$ patients have been treated and their responses observed, denote by $\hat{\rho}_{m,k}$ the estimated target allocation for treatment k.

The generalized **ERADE** algorithm with K treatments that we propose in our work assigns treatment $k \in [K]$ to patient $m \geq 1$ following the probability distribution $\{p_{m,k}\}_{k \in [K]}$ such that for all $k \in [K]$

$$\frac{N_{m,k}}{m} > \hat{\rho}_{m,k} \implies p_{m+1,k} \leq \alpha \hat{\rho}_{m,k} \tag{5.1}$$

and

$$\sum_{k=1}^{K} p_{m+1,k} = 1, \quad p_{m+1,k} \geq 0 \text{ for all } k \in [K] \tag{5.2}$$

where $\alpha \in [0, 1)$ controls the degree of randomization.

Remark 5.1. *Notice that the two treatments rule* (4.1) *can be expressed as* (5.1)*, the case when* $N_{m,k}/m > \rho_{m,k}$ *is directly satisfied. For the other case it is enough to satisfy* (5.2).

We keep using the same notations of section 2.1. We state the two following conditions, similar to the two arms case:

Condition A In the Bahadur-type representation (2.4) $\theta_k = E[\xi_{1,k}]$, and satisfying:

$$E\|\xi_{1,k}\|^2 < \infty, \quad k = 1, 2, \ldots, K.$$

Condition B The target allocation proportion function $\rho_k : \mathbb{R}^K \to (0, 1)$ is continuous and twice differentiable at Θ, with $\rho_k(\Theta) = v_k$

Example 5.1. *The allocation probabilities are updated according to*

$$p_{m+1,k} \;=\; \alpha \hat{\rho}_{m,k} \;+\; (1-\alpha)\,\frac{\delta_{m,k}}{\sum_{i=1}^{K} \delta_{m,i}}, \qquad \delta_{m,k} \;=\; \max\!\left(0,\; \hat{\rho}_{m,k} - \frac{N_{m,k}}{m}\right), \tag{5.3}$$

where $0 \leq \alpha < 1$

If for all j, $\delta_{m,j} = 0$, *then for every* $j \in \{1, \ldots, K\}$,

$$p_{m+1,j} \;=\; \hat{\rho}_{m,j}.$$

This rule combines two elements:

- *The term* $\hat{\rho}_{m,k}$ *represents the desired allocation proportion for treatment* k *at stage* m.
- *The imbalance correction* $\delta_{m,k}$ *measures whether treatment* k *is under-represented compared to its target. If* $\frac{N_{m,k}}{m} < \hat{\rho}_{m,k}$, *then* $\delta_{m,k} > 0$ *and the probability of assigning treatment* k *is increased. If* $\frac{N_{m,k}}{m} \geq \hat{\rho}_{m,k}$, *then* $\delta_{m,k} = 0$, *meaning no correction is applied.*

The parameter α *balances two forces:*

- *With* $\alpha = 1$, *the assignment is purely driven by the estimated target* $\hat{\rho}_{m,k}$.
- *With* $\alpha = 0$, *the rule focuses entirely on correcting imbalance by allocating more probability to under-represented treatments.*

Finally, if all treatments are exactly on target , then no correction is necessary.

5.2 Asymptotic Results

Let $\mathcal{F}_m = \sigma\left(X_1, \ldots, X_m, \xi_1, \ldots, \xi_m\right)$, $k \in [K]$ and an integer $n \geq 1$, we define

- $p_{m,k} = \mathbb{E}[X_{m,k} \mid \mathcal{F}_{m-1}]$
- $\Delta M_{m,k} = X_{m,k} - \mathbb{E}[X_{m,k} \mid \mathcal{F}_{m-1}] = X_{m,k} - p_{m,k}$
- $U_{n,k} = \sum_{m=1}^{n-1} \alpha\hat{\rho}_{m,k} + M_{n,k} - n\hat{\rho}_{n,k}$
- $Q_{n,k} = \sum_{m=1}^{n} X_{m,k}(\xi_{m,k} - \theta_k)$ and denote $\mathbf{Q_n} := (\mathbf{Q_{n,1}}, \ldots, \mathbf{Q_{n,K}})$
  Define the last time where $\frac{N_{m,k}}{m} \leq \hat{\rho}_{m,k}$

$$\ell_{n,k} := \max\left\{m \in \{Km_0 + 1, \ldots, n\} : \frac{N_{m,k}}{m} \leq \hat{\rho}_{m,k}\right\} \tag{5.4}$$

Lemma 5.1. *Some preliminary results that will be used later*

(i)

$$\Delta U_{n,k} = \alpha\hat{\rho}_{n-1,k} - p_{n,k} + \Delta\left(N_{n,k} - n\hat{\rho}_{n,k}\right) \tag{5.5}$$

(ii) If $\lim_{n\to+\infty} \hat{\rho}_n = v$ *a.s. then*

$$U_{n,k} - U_{\ell_{n,k},k} = (n - \ell_{n,k})[-(1-\alpha)v_k + o(1)] + M_{n,k} - M_{\ell_{n,k},k} + \ell_{n,k}(\hat{\rho}_{\ell_{n,k},k} - \hat{\rho}_{n,k}) \tag{5.6}$$

(iii)

$$N_{\ell_{n,k},k} - \ell_{n,k}\hat{\rho}_{\ell_{n,k},k} \leq Km_0 \tag{5.7}$$

(iv)

$$N_{n,k} - n\hat{\rho}_{n,k} \leq Km_0 + U_{n,k} - U_{\ell_{n,k},k} \tag{5.8}$$

Proof. (i) We have $\Delta M_{n,k} = \Delta N_{n,k} - p_{n,k} \implies \Delta U_{n,k} = \alpha \hat{\rho}_{n-1,k} - p_{n,k} + \Delta\left(N_{n,k} - n\hat{\rho}_{n,k}\right)$

(ii)

$$\begin{aligned}
U_{n,k} - U_{\ell_{n,k},k} &= \sum_{m=\ell_{n,k}}^{n-1} \alpha \hat{\rho}_{m,k} - n\hat{\rho}_{n,k} + \ell_{n,k}\hat{\rho}_{\ell_{n,k},k} + M_{n,k} - M_{\ell_{n,k},k} \\
&= \sum_{m=\ell_{n,k}}^{n-1} \alpha(\hat{\rho}_{m,k} - v_k) + \alpha(n - \ell_{n,k})v_k - (n - \ell_{n,k})\hat{\rho}_{n,k} + \ell_{n,k}(\hat{\rho}_{\ell_{n,k},k} - \hat{\rho}_{n,k}) + M_{n,k} - M_{\ell_{n,k},k} \\
&= -(1-\alpha)(n - \ell_{n,k})v_k + \sum_{m=\ell_{n,k}}^{n-1} \alpha(\hat{\rho}_{m,k} - v_k) - (n - \ell_{n,k})(\hat{\rho}_{n,k} - v_k) + M_{n,k} - M_{\ell_{n,k},k} \\
&\quad + \ell_{n,k}(\hat{\rho}_{\ell_{n,k},k} - \hat{\rho}_{n,k}) \\
&= (n - \ell_{n,k})[-(1-\alpha)v_k + o(1)] + M_{n,k} - M_{\ell_{n,k},k} + \ell_{n,k}(\hat{\rho}_{\ell_{n,k},k} - \hat{\rho}_{n,k})
\end{aligned}$$

(iii) It is easy to check that if $\ell_{n,k} < Km_0 + 1$ then

$$N_{\ell_{n,k},k} - \ell_{n,k}\hat{\rho}_{\ell_{n,k},k} \leq Km_0$$

and if $\ell_{n,k} \geq Km_0 + 1$ then

$$N_{\ell_{n,k},k} - \ell_{n,k}\hat{\rho}_{\ell_{n,k},k} \leq 0$$

so by combining the two inequalities we obtain that $N_{\ell_{n,k},k} - \ell_{n,k}\hat{\rho}_{\ell_{n,k},k} \leq Km_0$

(iv) With the definition of $\ell_{n,k}$ in (5.4) we have

$$\forall \ell_{n,k} + 1 \leq m \leq n : \frac{N_{m,k}}{m} > \hat{\rho}_{m,k}$$

And in formula (5.5), we have for $m \geq \ell_{n,k} + 1$

$$\Delta U_{n,k} \geq \Delta\left(N_{n,k} - n\hat{\rho}_{n,k}\right) \tag{5.9}$$

so summing the formula (5.9) from $\ell_{n,k} + 1$ to n yields

$$N_{n,k} - n\hat{\rho}_{n,k} = N_{\ell_{n,k},k} - \ell_{n,k}\hat{\rho}_{\ell_{n,k},k} + U_{n,k} - U_{\ell_{n,k},k}$$

Then using (5.7) we obtain (5.8)

□

Lemma 5.2. *Under conditions **A** and **B** we have as $n \to +\infty$:*

$$\lim_{n\to\infty} \frac{N_{n,k}}{n} = v_k \text{ a.s.} \tag{5.10}$$

$$\hat{\Theta}_n - \Theta = O\left(\sqrt{\frac{\log\log n}{n}}\right) \text{ a.s.} \tag{5.11}$$

An immediate result is the following

$$n\left(\hat{\rho}_{n,k} - v\right) = O\left(\sqrt{n\,\text{loglog}(n)}\right) \text{ a.s.} \tag{5.12}$$

Proof. By Lemma 3.2 in section 3, we proved that $\hat{\rho}_{n,k}$ has finite limit, so with the same argument; by continuity of ρ we have

$$\exists u \in (0,1)^K : \quad \hat{\rho}_n \to u \tag{5.13}$$

We also have by Lemma C.1 that $M_{n,k} = o(n)$ a.s. so

$$\begin{aligned}
\frac{U_{n,k}}{n} &= \frac{1}{n}\sum_{m=0}^{n-1} \alpha\hat{\rho}_{m,k} - \hat{\rho}_{n,k} + o(1) \\
\implies \lim_{n\to\infty} \frac{U_{n,k}}{n} &= -(1-\alpha)u_k \quad \text{(by convergence (5.13) and Césaro lemma)} \\
\implies U_{n,k} &\sim -n(1-\alpha)u_k \quad \text{a.s.}
\end{aligned}$$

Using the Lemma B.1 we obtain that:

$$\left(\frac{Km_0}{n} + \frac{U_{n,k} - U_{\ell_{n,k},k}}{n}\right)^+ \to 0$$

So in all cases:

$$\left(\frac{N_{n,k}}{n} - \hat{\rho}_{n,k}\right)^+ \to 0$$

And we have:

$$\begin{aligned}
\left(\hat{\rho}_{n,k} - \frac{N_{n,k}}{n}\right)^+ &= \left(1 - \frac{N_{n,k}}{n} - (1-\hat{\rho}_{n,k})\right)^+ \\
&= \left(\sum_{j=1;j\neq k}^{K} \frac{N_{n,j}}{n} - \hat{\rho}_{n,j}\right)^+ \\
&\leq \sum_{j=1;j\neq k}^{K} \left(\frac{N_{n,j}}{n} - \hat{\rho}_{n,j}\right)^+ \to 0
\end{aligned}$$

So we obtain that $\frac{N_{n,k}}{n} - \hat{\rho}_{n,k} \to 0$ and then $\lim_{n\to\infty} \frac{N_{n,k}}{n} = \lim_{n\to\infty} \hat{\rho}_{n,k} = u_k \in (0,1)$

By continuity of ρ we obtain that $\hat{\rho}_n = \rho(\hat{\Theta}_n) \to \rho(\Theta) = v \in (0,1)^K$ then

$$\lim_{n\to\infty} \frac{N_{n,k}}{n} = \lim_{n\to\infty} \hat{\rho}_{n,k} = v_k$$

An immediate result is that $\forall k \in [K], N_{n,k} \to \infty$ a.s. then by Zhang (2004) Lemma A.3 Hu and Zhang (2004) we get:

$$\hat{\Theta}_n - \Theta = O\left(\sqrt{\frac{\log\log n}{n}}\right) \quad \text{a.s.}$$

$\square$

Lemma 5.3. *Let $(\ell_n)_{n\geq 1}$ a non-decreasing sequence of positive integers such that $\ell_n \leq n$. Under conditions A and B*

$$|\ell_n(\hat{\rho}_n - \hat{\rho}_{\ell_n})| \leq o(1)(n-\ell_n) + C\|Q_n - Q_{\ell_n}\|$$

Proof. The second order Taylor expansion of ρ and Lemma C.4 give

$$\hat{\rho}_n - v = (\hat{\Theta}_n - \Theta)\frac{\partial \rho}{\partial y}\Big|_{\Theta} + O(\|\hat{\Theta}_n - \Theta\|^2) \tag{5.14}$$

$$= (\hat{\Theta}_n - \Theta)\frac{\partial \rho}{\partial y}\Big|_{\Theta} + O_p\left(\frac{1}{n}\right) \tag{5.15}$$

Similarly we obtain:

$$\hat{\rho}_{\ell_n} - v = (\hat{\Theta}_{\ell_n} - \Theta)\frac{\partial \rho}{\partial y}\Big|_{\Theta} + O_p\left(\frac{1}{\ell_n}\right)$$

Combining the two previous equations gives:

$$\begin{aligned}\ell_n(\hat{\rho}_{\ell_n} - \hat{\rho}_n) &= \ell_n(\hat{\Theta}_{\ell_n} - \hat{\Theta}_n)\frac{\partial \rho}{\partial y}\Big|_{\Theta} + \left(O_p\left(\frac{\ell_n}{n}\right) + O_p(1)\right) \\ &\implies |\ell_n(\hat{\rho}_{\ell_n} - \hat{\rho}_n)| \le C\ell_n\|\hat{\Theta}_{\ell_n} - \hat{\Theta}_n\| + O_p(1)\end{aligned}$$

And we have:

$$\begin{aligned}\hat{\theta}_{n,k} - \hat{\theta}_{\ell_n,k} &= \frac{1}{N_{n,k}}\sum_{j=1}^{n} X_{j,k}(\xi_{j,k} - \theta_k) - \frac{1}{N_{\ell_n,k}}\sum_{j=1}^{\ell_n} X_{j,k}(\xi_{j,k} - \theta_k) \\ &= \frac{1}{N_{n,k}}\sum_{j=\ell_n+1}^{n} X_{j,k}(\xi_{j,k} - \theta_k) + \left(\frac{1}{N_{n,k}} - \frac{1}{N_{\ell_n,k}}\right)\sum_{j=1}^{\ell_n} X_{j,k}(\xi_{j,k} - \theta_k) \\ &= \frac{1}{N_{n,k}}\sum_{j=\ell_n+1}^{n} X_{j,k}(\xi_{j,k} - \theta_k) + \frac{N_{\ell_n,k} - N_{n,k}}{N_{n,k}} \cdot \frac{1}{N_{\ell_n,k}}\sum_{j=1}^{\ell_n} X_{j,k}(\xi_{j,k} - \theta_k) \\ &\quad + o\left(N_{n,k}^{-1/2}\right) + o\left(N_{\ell_n,k}^{-1/2}\right)\end{aligned}$$

So

$$\begin{aligned}|\ell_n(\hat{\rho}_{\ell_n} - \hat{\rho}_n)| \le{}& C\sum_{k=1}^{K}\frac{\ell_n}{N_{n,k}}\left\|\sum_{j=\ell_n+1}^{n} X_{j,k}(\xi_{j,k} - \theta_k)\right\| \\ &+ C\sum_{k=1}^{K}\frac{\ell_n|N_{\ell_n,k} - N_{n,k}|}{N_{n,k}}\left\|\frac{1}{N_{\ell_n,k}}\sum_{j=1}^{\ell_n} X_{j,k}(\xi_{j,k} - \theta_k)\right\| \\ &+ \ell_n \cdot o\left(N_{n,k}^{-1/2}\right) + \ell_n \cdot o\left(N_{\ell_n,k}^{-1/2}\right)\end{aligned}$$

And for some two sequences a_n and b_n converging to 0 as $n \to +\infty$

$$\begin{aligned}\ell_n \cdot o\left(N_{n,k}^{-1/2}\right) + \ell_n \cdot o\left(N_{\ell_n,k}^{-1/2}\right) &= \ell_n\, N_{n,k}^{-1/2} a_n + \ell_n\, N_{\ell_n,k}^{-1/2} b_n \\ &= \sqrt{n}\left(\frac{\ell_n}{n} \cdot \left(\frac{N_{n,k}}{n}\right)^{-1/2} a_n + \sqrt{\frac{\ell_n}{n}}\left(\frac{N_{\ell_n,k}}{\ell_n}\right)^{-1/2} b_n\right) \\ &= o(\sqrt{n}) = o_p(\sqrt{n})\end{aligned}$$

As $\lim\limits_{n\to+\infty} \frac{\ell_n}{n}\left(\frac{N_{n,k}}{n}\right)^{-1/2} a_n + \sqrt{\frac{\ell_n}{n}}\left(\frac{N_{\ell_n,k}}{\ell_n}\right)^{-1/2} b_n = 0$

We have:

$$\begin{aligned}C\sum_{k=1}^{K}\frac{\ell_n}{N_{n,k}}\left\|\sum_{m=\ell_n+1}^{n} X_{m,k}(\xi_{m,k} - \theta_k)\right\| &= C\sum_{k=1}^{K}\frac{\ell_n}{n}\frac{n}{N_{n,k}}\|Q_{n,k} - Q_{\ell_n,k}\| \\ &\le C\sum_{k=1}^{K}\|Q_{n,k} - Q_{\ell_n,k}\| \quad (\text{since } \ell_n \le n \text{ and } n/N_{n,k} \to v_k^{-1}) \\ &= C\|Q_n - Q_{\ell_n}\|\end{aligned}$$

On the second part of the inequality we have:

$$
\begin{aligned}
C\sum_{k=1}^{K}\frac{\ell_n|N_{\ell_n,k}-N_{n,k}|}{N_{n,k}}\left\|\frac{1}{N_{\ell_n,k}}\sum_{m=1}^{\ell_n}X_{m,k}(\xi_{m,k}-\theta_k)\right\| &= C\sum_{k=1}^{K}\frac{\ell_n|N_{\ell_n,k}-N_{n,k}|}{N_{\ell_n,k}N_{n,k}}\left\|\sum_{m=1}^{\ell_n}X_{m,k}(\xi_{m,k}-\theta_k)\right\| \\
&= v_k^{-1}\left(1-\frac{\ell_n}{n}\right)\left\|\sum_{m=1}^{\ell_n}X_{m,k}(\xi_{m,k}-\theta_k)\right\| \\
&= v_k^{-1}(n-\ell_n)\frac{1}{n}\left\|\sum_{m=1}^{\ell_n}X_{m,k}(\xi_{m,k}-\theta_k)\right\|
\end{aligned}
$$

If $(\ell_n)_n$ is bounded, then $\frac{\ell_n}{n}\frac{1}{\ell_n}\left\|\sum_{m=1}^{\ell_n}X_{m,k}(\xi_{m,k}-\theta_k)\right\|=o(1)$. Otherwise if $\ell_n\to+\infty$ then we can use the argument of Doob (1936) and conclude that $\frac{1}{\ell_n}\left\|\sum_{m=1}^{\ell_n}X_{m,k}(\xi_{m,k}-\theta_k)\right\|=o(1)$. So in both cases we have

$$
\begin{aligned}
\frac{1}{n}\left\|\sum_{m=1}^{\ell_n}X_{m,k}(\xi_{m,k}-\theta_k)\right\| &\leq \frac{1}{\ell_n}\left\|\sum_{m=1}^{\ell_n}X_{m,k}(\xi_{m,k}-\theta_k)\right\| \\
&= o(1)
\end{aligned}
$$

So

$$
|\ell_n(\hat{\rho}_n-\hat{\rho}_{\ell_n})|\leq o(1)(n-\ell_n)+C\|Q_n-Q_{\ell_n}\|
$$

□

Lemma 5.4. *Under conditions **A** and **B** we have fort every $k\in[K]$ as $n\to+\infty$*

$$
U_{n,k}-U_{\ell_{n,k},k}\leq o_p(\sqrt{n}) \tag{5.16}
$$

$$
U_{n,k}-U_{\ell_{n,k},k}\leq O(\sqrt{n\log\log(n)})\quad a.s. \tag{5.17}
$$

Proof. From (5.6) in Lemma 5.1 we have

$$
U_{n,k}-U_{\ell_{n,k},k}=(n-\ell_{n,k})[-(1-\alpha)v_k+o(1)]+M_{n,k}-M_{\ell_{n,k},k}+\ell_{n,k}(\hat{\rho}_{\ell_{n,k},k}-\hat{\rho}_{n,k}) \tag{5.18}
$$

From Lemma 5.3 we have

$$
|\ell_{n,k}(\hat{\rho}_{\ell_{n,k},k}-\hat{\rho}_{n,k})|\leq o(1)(n-\ell_{n,k})+C\|Q_n-Q_{\ell_{n,k}}\|+o_p(\sqrt{n}) \tag{5.19}
$$

We also have from Lemma B.3 that:

$$
\|Q_n-Q_{\ell_{n,k}}\|\leq(n-\ell_{n,k})\max_{L\leq m<n}\frac{\|Q_n-Q_{n-m}\|}{m}+\max_{m<L}\|Q_n-Q_{n-m}\| \tag{5.20}
$$

Let $(L_n)_{n\geq1}$ a positive sequence such that $L_n\to\infty$ and $L_n=o(n)$, so from Lemma B.2 we have:

$$
\mathbb{E}\left[\max_{m\leq L_n}\|Q_n-Q_{n-m}\|\right]\leq 2\sqrt{L_nC_0}\quad\text{and}\quad\mathbb{E}\left[\max_{L_n\leq m<n}\frac{\|Q_n-Q_{n-m}\|}{m}\right]\leq 3\sqrt{\frac{C_0}{L_n}}
$$

Applying Lemma C.5 we obtain,

$$
\max_{L_n\leq m<n}\frac{\|Q_m-Q_{n-m}\|}{m}=O_p\left(\frac{1}{\sqrt{L_n}}\right)=o_p(1)\quad\text{and}\quad\max_{L_n\leq m<n}\|Q_m-Q_{n-m}\|=O_p(\sqrt{L_n})\leq o_p(\sqrt{n})
$$

So applying this in (5.20) we get,

$$
\|Q_n-Q_{\ell_{n,k}}\|\leq o_p(1)(n-\ell_{n,k})+o_p(\sqrt{n})
$$

We use exactly the same steps to obtain the same result on the martingale $\{M_n\}_{n\geq 1}$

$$|M_{n,k} - M_{\ell_{n,k},k}| \leq o_p(1)(n - \ell_{n,k}) + o_p(\sqrt{n})$$

Therefore, using the last two inequalities together with (5.18) and (5.19) we obtain that

$$\begin{aligned}
U_{n,k} - U_{\ell_{n,k},k} &\leq (n-\ell_{n,k})[-(1-\alpha)v_k + o(1)] + |M_{n,k} - M_{\ell_{n,k},k}| + \ell_{n,k}|\hat{\rho}_{\ell_{n,k},k} - \hat{\rho}_{n,k}| \\
&\leq (n-\ell_{n,k})[-(1-\alpha)v_k + o(1)] + |M_{n,k} - M_{\ell_{n,k},k}| + o(1)(n-\ell_{n,k}) + C\|Q_n - Q_{\ell_{n,k}}\| + o_p(\sqrt{n}) \\
&\leq (n-\ell_{n,k})[-(1-\alpha)v_k + o(1)] + o_p(1)\,(n-\ell_{n,k}) + o_p(\sqrt{n}) + o(1)\,(n-\ell_{n,k}) \\
&= (n-\ell_{n,k})\underbrace{[-(1-\alpha)v_k + o(1)]}_{\leq 0} + o_p(\sqrt{n}) \\
&\leq o_p(\sqrt{n}) \quad \text{a.s.}
\end{aligned}$$

We proved in Lemma 5.2 that $n(\hat{\rho}_n - v) = o\left(\sqrt{n\log\log n}\right)$ a.s. which also gives that

$$\begin{aligned}
\ell_{n,k}(\hat{\rho}_{\ell_{n,k}} - \hat{\rho}_{n,k}) &= o\left(\sqrt{n\log\log n}\right) + o\left(\sqrt{\ell_{n,k}\log\log \ell_{n,k}}\right) \\
&= o\left(\sqrt{n\log\log n}\right) \quad \text{a.s.}
\end{aligned}$$

Also, $\{M_{n,k}\}_{n\geq 1}$ is a martingale, so by LIL for martingales (Theorem C.2) we have $M_{n,k} = o\left(\sqrt{n\,\text{loglog}\, n}\right)$ then

$$\begin{aligned}
M_{n,k} - M_{\ell_{n,k},k} &= o\left(\sqrt{n\,\text{loglog}\, n}\right) + o\left(\sqrt{\ell_{n,k}\,\text{loglog}\,\ell_{n,k}}\right) \\
&= o\left(\sqrt{n\,\text{loglog}\, n}\right) \quad \text{a.s.}
\end{aligned}$$

back to (5.18), as $1-\alpha > 0$ we have

$$\begin{aligned}
U_{n,k} - U_{\ell_{n,k},k} &\leq M_{n,k} - M_{\ell_{n,k},k} + \ell_{n,k}(\hat{\rho}_{\ell_{n,k},k} - \hat{\rho}_{n,k}) \\
&\leq O(\sqrt{n\,\text{loglog}\, n}) \quad \text{a.s.}
\end{aligned}$$

□

We now state the main asymptotic properties of the **ERADE** extension with K treatments with the notations $G = \nabla(\rho)\big|_{\Theta}$; $V_k = \mathbb{V}ar[\xi_{1,k}]$ for $k \in [K]$, and

$$V = diag\left(\frac{1}{v_1}V_1, \ldots, \frac{1}{v_K}V_K\right) \text{ and } \Omega = \begin{pmatrix} GVG^T & GVG^T \\ GVG^T & GVG^T \end{pmatrix}$$

Theorem 5.1. *Under conditions **A** and **B** we have*

(i) For every $k \in [K]$ as $n \to \infty$,

$$|N_{n,k} - n\hat{\rho}_{n,k}| = o_P(\sqrt{n}) \tag{5.21}$$
$$|N_{n,k} - n\hat{\rho}_{n,k}| = O(\sqrt{n\log\log n}) \quad a.s. \tag{5.22}$$
$$N_{n,k} - nv_k = O(\sqrt{n\log\log n}) \quad a.s. \tag{5.23}$$

(ii) The asymptotic normality results:

$$\sqrt{n}(\hat{\Theta}_n - \Theta) \xrightarrow{\mathcal{D}} N(0, V), \tag{5.24}$$
$$\begin{pmatrix} \sqrt{n}\left(\frac{N_n}{n} - v\right) \\ \sqrt{n}(\hat{\rho}_n - v) \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Omega). \tag{5.25}$$

Proof. (i) Using Lemma 5.4, inequality (5.8) from Lemma 5.1 give that

$$\forall k \in [K] : N_{n,k} - n\hat{\rho}_{n,k} \leq o_p(\sqrt{n}) \quad \text{and} \quad N_{n,k} - n\hat{\rho}_{n,k} \leq O(\sqrt{n \log\log n}) \quad \text{a.s.} \tag{5.26}$$

And we have that

$$\begin{aligned} n\hat{\rho}_{n,k} - N_{n,k} &= n - N_{n,k} - n(1 - \hat{\rho}_{n,k}) && (5.27)\\ &= \sum_{j=1;j\neq k}^{K} (N_{n,j} - n\hat{\rho}_{n,j}) && (5.28)\\ &\leq \sum_{j=1;j\neq k}^{K} o_p(\sqrt{n}) \quad \text{a.s.} && (5.29)\\ &= o_p(\sqrt{n}) \quad \text{a.s.} && (5.30)\end{aligned}$$

We apply the same steps for the second inequality and get that $n\hat{\rho}_{n,k} - N_{n,k} \leq O\left(\sqrt{n\,\mathrm{loglog}(n)}\right)$ a.s., so with (5.26) we obtain (5.21) and (5.22),

$$|N_{n,k} - n\hat{\rho}_{n,k}| = o_p(\sqrt{n}) \quad \text{and} \quad |N_{n,k} - n\hat{\rho}_{n,k}| = O(\sqrt{n\,\mathrm{loglog}\, n}) \quad \text{a.s.}$$

now using (5.12) and (5.18) we have

$$\begin{aligned} N_{n,k} - nv_k &= N_{n,k} - n\hat{\rho}_{n,k} + n(\hat{\rho}_{n,k} - v_k)\\ &= O(\sqrt{n \log\log n}) \quad a.s. \end{aligned}$$

which proves (5.23).

(ii) For the asymptotic normality results, using Lemma 1 from Hu et al. (2006) gives that

$$\sqrt{n}(\hat{\Theta}_n - \Theta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, V) \tag{5.31}$$

And we have using (5.15)

$$\begin{aligned} \hat{\rho}_n - v &= (\hat{\Theta}_n - \Theta)\, G + O_p\left(\frac{1}{n}\right)\\ \implies \sqrt{n}(\hat{\rho}_n - v) &= \sqrt{n}(\hat{\Theta}_n - \Theta) G + O_p\left(\frac{1}{\sqrt{n}}\right) \end{aligned}$$

From $O_p\left(\frac{1}{\sqrt{n}}\right) \xrightarrow{\mathbb{P}} 0$ and (5.31), we can apply Slutsky theorem giving

$$\sqrt{n}(\hat{\rho}_n - v) \xrightarrow{\mathcal{D}} \mathcal{N}(0, GVG^T) \tag{5.32}$$

And we have

$$\begin{aligned} \sqrt{n}\left(\frac{N_n}{n} - v\right) &= \sqrt{n}\left(\frac{N_n}{n} - \hat{\rho}_n\right) + \sqrt{n}(\hat{\rho}_n - \nu)\\ &= \sqrt{n}(\hat{\rho}_n - \nu) + o_p(1) \end{aligned}$$

Again with (5.32) we can apply Slutsky theorem:

$$\sqrt{n}\left(\frac{N_n}{n} - v\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, GVG^T)$$

Noting $W_n := \sqrt{n}\left(\frac{N_n}{n} - \nu\right)$ and $Y_n := \sqrt{n}(\hat{\rho}_n - \nu)$, then

$$\begin{pmatrix} W_n \\ Y_n \end{pmatrix} = \begin{pmatrix} Y_n \\ Y_n \end{pmatrix} + \underbrace{\begin{pmatrix} o_p(1) \\ 0 \end{pmatrix}}_{\xrightarrow{\mathbb{P}} 0}$$

And as we have

$$\begin{pmatrix} Y_n \\ Y_n \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \begin{pmatrix} GVG^T & GVG^T \\ GVG^T & GVG^T \end{pmatrix}\right),$$

we can use Slutsky again and get

$$\begin{pmatrix} W_n \\ Y_n \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Omega).$$

$\square$

Theorem 5.2. *Under conditions* $\boldsymbol{A}$ *and* $\boldsymbol{B}$*, if* $\mathbb{V}ar\,(\xi_{1,k}) = I_k(\theta_k)^{-1}$ *for all* $k \in [K]$*, then the asymptotic variance of* $\frac{N_n}{\sqrt{n}}$ *reaches the lower bound specified by Hu et al. (2006) as*

$$G\,I(\Theta)^{-1}\,G^T$$

where $I(\Theta) = diag\,(v_1 I_1(\theta_1), \ldots, v_K I_K(\theta_K))$ *and* $I_k(\theta_k)$ *is the fisher information for a single observation on treatment* $k \in [K]$*.*

Remark 5.2. *For DBCD design, Hu and Zhang (2004) proved the asymptotic normality of the allocation proportions when using the allocation function* g^γ *defined in* (3.1)*, given by*

$$\sqrt{n}\left(\frac{N_n}{n} - v\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Omega_\gamma) \quad \textit{where} \quad \Omega_\gamma = \frac{1}{1+2\gamma}\Sigma_1 + \frac{2(1+\gamma)}{1+2\gamma} G\,I(\Theta)^{-1}\,G^T$$

Notice that

$$\Omega_\gamma - G\,I(\Theta)^{-1}\,G^T = \frac{1}{1+2\gamma}\left(\Sigma_1 + G\,I(\Theta)^{-1}\,G^T\right) \succeq 0$$

which means that DBCD does not reach the variance lower bound asymptotically, the design approaches to this lower bound as γ *increases, reaching it when* $\gamma \to +\infty$*.*

This behaviour can be explained by the fact that the allocation function g^γ *used in Hu and Zhang (2004) is the same as an ERADE allocation function when* $\gamma \to +\infty$*. In fact considering* g^γ *with a cap* $L > 0$

$$g_j^\gamma(x, y) = \frac{y_j\,(y_j/x_j)^\gamma \wedge L}{\sum_{i=1}^K y_i\,(y_j/x_j)^\gamma \wedge L}; \quad \textit{for } j \in [K]$$

taking the example of $K = 2$*, we obtain using the ERADE allocation function formulation given in* (4.2)

$$\lim_{\gamma \to +\infty} g^\gamma(x, y) = g(x, y; 0) = y\,\mathbf{1}_{x=y} + \mathbf{1}_{x<y} \tag{5.33}$$

For $K > 2$*, consider the set* $S = \left\{j; \quad \frac{y_j}{x_j} = \max_{i\in[K]} \frac{y_i}{x_i}\right\}$*, we have*

$$\lim_{\gamma \to +\infty} g_j^\gamma(x, y) = \begin{cases} \frac{1}{|S|}, & \textit{if } j \in S, \\ 0, & \textit{if } j \notin S, \end{cases} \tag{5.34}$$

which is also a special case of the generalized ERADE to k *treatments taking* $\alpha = 0$

5.3 Numerical Simulation

For the numerical experiments, we fixed the following parameters:

- Initial sample size per treatment: $m_0 = 2$,
- Number of treatments: $K = 4$,

- Total number of patients: $n = 10^4$,
- Initial parameter value: $\theta_0 = 0.5$.

The allocation target $\rho(\Theta)$ was chosen according to the following

$$\hat{\rho}_k \;=\; \frac{\dfrac{1}{1-\hat{\theta}_k}}{\sum_{j=1}^{K} \dfrac{1}{1-\hat{\theta}_j}}, \qquad k = 1, \ldots, K.$$

Intuitively, treatments with smaller failure probabilities $(1 - \hat{\theta}_k)$ receive larger weights, so that the design tends to allocate more patients to treatments estimated to be more effective.

We use the allocation rule described in (5.3). We present simulation results for three different choices of the randomization parameter $\alpha \in \{0, 0.3, 0.9\}$. The effect of α is to control the balance between determinism and randomness in the allocation procedure.

Case $\alpha = 0$**.** As shown in Figure 1, the procedure is almost deterministic. The allocation proportions $\frac{N_n}{n}$ and $\hat{\rho}_n$ converge rapidly to the target v with very small fluctuations. The deviation $\left|\frac{N_n}{n} - \hat{\rho}_n\right|$ decreases essentially at the $1/n$ rate, as predicted. Scaled quantities stabilize tightly around zero, confirming the theoretical asymptotic results. This case shows minimal variability and strong finite-sample performance.

Case $\alpha = 0.3$**.** Figure 2 shows that introducing moderate randomization increases the fluctuations of $\frac{N_n}{n}$ and $\hat{\rho}_n$, but the convergence remains clear. The deviation $\left|\frac{N_n}{n} - \hat{\rho}_n\right|$ follows the $1/n$ order, though with slightly higher noise. The scaled quantities stabilize with mild oscillations. This value of α achieves a good trade-off between maintaining randomization and preserving efficiency.

Case $\alpha = 0.9$**.** For high randomization, shown in Figure 3, variability dominates the allocation. The convergence of $\frac{N_n}{n}$ and $\hat{\rho}_n$ becomes noisy and slower to stabilize. The deviation $\left|\frac{N_n}{n} - \hat{\rho}_n\right|$ decays more slowly, and scaled quantities fluctuate strongly instead of collapsing neatly to zero. While the asymptotic properties are valid, finite-sample behavior is poor and unstable.

5.4 Example: Asymptotic Properties for Track and Stop Algorithm

The Track and Stop Strategy has been introduced by Garivier and Kaufmann (2016). As stated in the paper, the first approach to aligning the proportions $\omega^*(\mu)$ is to follow the plug-in estimates $\omega^*(\hat{\mu}(t))$. However, in bandit problems, relying directly on plug-in estimates can be risky: a poor initial estimate may cause an arm to be discarded too early, preventing the collection of new data that could have corrected the initial mistake. As observed both in theory and simulations, such a naive plug-in strategy may indeed fail. A straightforward remedy is to enforce sufficient exploration of every arm, so as to guarantee a (sufficiently rapid) convergence of $\hat{\mu}(t)$.

Let $\mathcal{A} = \{1, \ldots, K\}$ be the set of arms, and for each $a \in \mathcal{A}$ let $N_a(t)$ be the number of pulls of arm a up to time t, and $\widehat{\mu}_a(t)$ its empirical mean. Denote $\widehat{\mu}(t) = (\widehat{\mu}_1(t), \ldots, \widehat{\mu}_K(t))$ and let $w^\star(\cdot) \in \Sigma_K$ be the vector of *optimal sampling proportions* returned by the characteristic-time optimization (computed at the current estimate).

At round t, define the under-sampled set

$$U_t := \left\{ a \in \mathcal{A} \;:\; N_a(t) < \sqrt{t} - \tfrac{K}{2} \right\}.$$

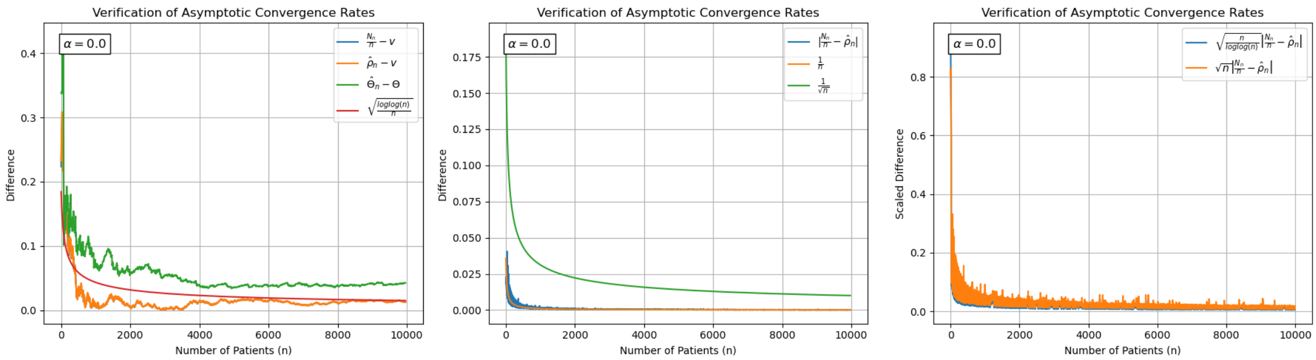

Figure 1: Verification of asymptotic convergence rates for $\alpha = 0$

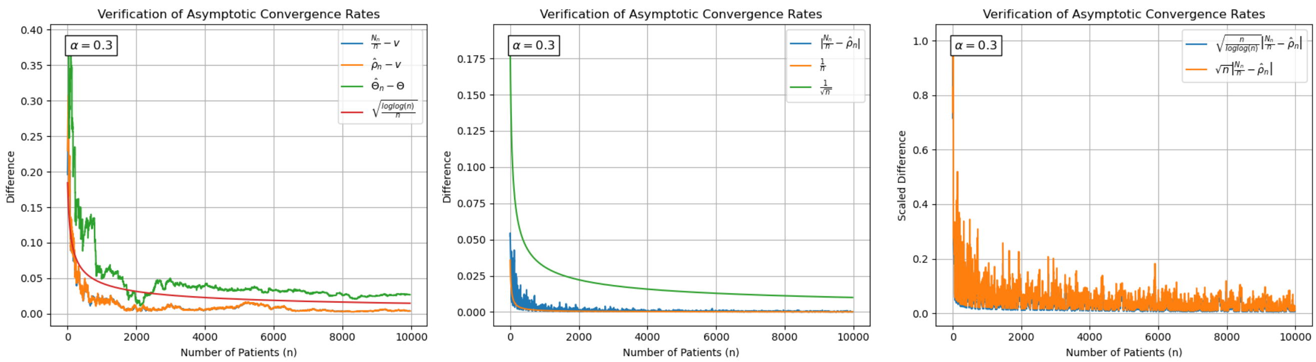

Figure 2: Verification of asymptotic convergence rates for $\alpha = 0.3$

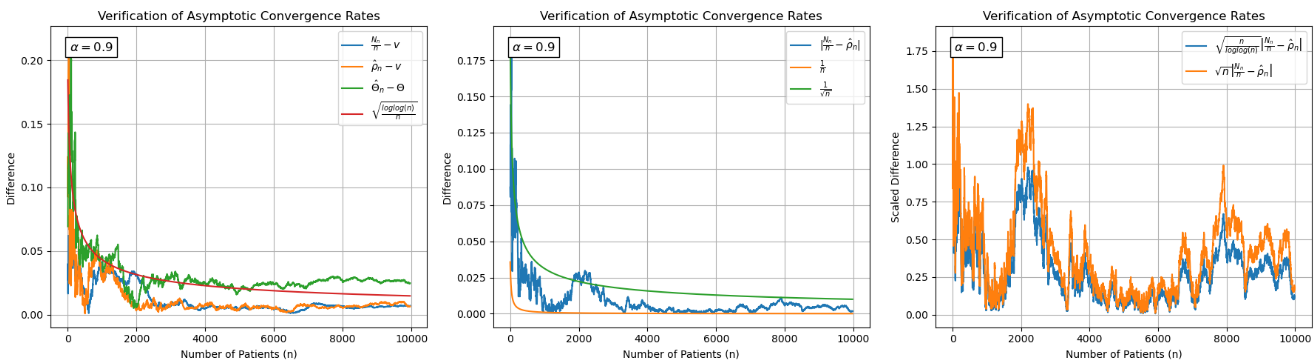

Figure 3: Verification of asymptotic convergence rates for $\alpha = 0.9$

Given $\widehat{\mu}(t)$ and $N_1(t), \ldots, N_K(t)$, the called **D-Tracking** rule consists in choosing the next arm A_{t+1} by

$$A_{t+1} \in \begin{cases} \arg\min\limits_{a \in U_t} N_a(t), & \text{if } U_t \neq \emptyset \quad \text{(Forced Exploration)}, \\ \arg\max\limits_{a \in \mathcal{A}} \Big\{ t\, w_a^\star\big(\widehat{\mu}(t)\big) - N_a(t) \Big\}, & \text{if } U_t = \emptyset \quad \text{(Direct Tracking)}. \end{cases} \tag{5.35}$$

Remark 5.3. *In relation to the clinical trial context, the function* $\omega_k^*\left(\hat{\mu}(m)\right)$ *will play the same role of the allocation target* $\hat{\rho}_{m,k}$.

In our setting as in (5.1), we can propose a **D-Tracking**-like algorithm; we will consider the **D-Tracking** as in (5.35) but without the *Forced Exploration* phase. Using our notation, the proposed algorithm will follow the allocation rule

$$p_{m+1,k} = \begin{cases} 0 & , \quad \text{if } \frac{N_{m,k}}{m} > \hat{\rho}_{m,k} \\ \mathbb{1}_{k=k^\star} & , \quad \text{else} \end{cases} \tag{5.36}$$

where $k^\star = \arg\max\limits_{a\in[K]} \Big\{ m\, \hat{\rho}_{m,a} - N_{m,a} \Big\}$.

Remark 5.4. *In case of* $k^\star$ *not unique, we can select uniformly one of them and consider it for* (5.36).

Remark 5.5. *The rule* (5.36) *is the same as the DBCD allocation rule using the allocation function* g^γ *when* $\gamma \to +\infty$ *as seen in* (5.34).

We can see that the proposed allocation for **D-Tracking**-like algorithm follows the scheme of generalized **ERADE** (5.1) with $\alpha = 0$. As a result we conclude all the asymptotic and CLT results given in Theorem 5.1 for this type of Best Arm Identification algorithms.

What about the real D-Tracking? Designing **D-Tracking** in our framework is not straightforward, but we can present a different type of allocation rule than (5.1) that can contain **D-Tracking**. For $k \in [K]$ and $m \geq 1$ consider

$$p_{m+1,k} = \alpha \hat{\rho}_{m,k} + (1-\alpha) \frac{\delta_{m,k}}{\sum_{j=1}^K \delta_{m,j}}; \quad \alpha \in [0,1)$$

The **D-tracking** algorithm can be expressed equivalently in our setup by taking $\alpha = 0$ and

$$\delta_{m,k}\left(\frac{N_{m,k}}{m}, \hat{\rho}_{m,k}\right) = \begin{cases} 1, & \text{if } FE_{m,k} \\ 1, & \text{if } DT_{m,k} \\ 0, & \text{else} \end{cases}$$

Where

$$\begin{aligned} FE_{m,k} &:= \{k = arg\min_{j\in U_m} N_{m,k} \text{ and } U_m \neq \emptyset\} \\ DT_{m,k} &:= \{k = arg\max_{1\leq j\leq K} \{m\hat{\rho}_{m,j} - N_{m,j}\} \text{ and } U_m \neq \emptyset\} \\ U_m &:= \left\{ j \in [K] : \quad N_{m,j} < \sqrt{m} - \frac{K}{2} \right\} \end{aligned}$$

We can prove the convergence of proportions $\frac{N_n}{n}$ to allocation targets v in this setting.

Sketch of proof. We need to redefine the stopping time (5.4) for this different setting, consider

$$\ell_{n,k} = \max\Big\{1 \leq m \leq n : \ FE_{k,m} \text{ or } DT_{k,m}\Big\}$$

With the previous notations we have

$$U_{n,k} - U_{\ell_{n,k},k} = -\underbrace{\sum_{m=\ell_{n,k}+1}^{n} p_{m,k}}_{=0} + N_{n,k} - n\hat{\rho}_{n,k} - N_{\ell_{n,k},k} + \ell_{n,k}\hat{\rho}_{\ell_{n,k},k} \tag{5.37}$$

Then

$$N_{n,k} - n\hat{\rho}_{n,k} = U_{n,k} - U_{\ell_{n,k},k} + N_{\ell_{n,k},k} - \ell_{n,k}\hat{\rho}_{\ell_{n,k},k} \tag{5.38}$$

In the event $FE_{m,k}$, we have:

$$N_{\ell_{n,k},k} - \ell_{n,k}\hat{\rho}_{\ell_{n,k},k} \leq N_{\ell_{n,k},k} \leq \sqrt{\ell_{n,k}} - \frac{K}{2} \tag{5.39}$$

And for the event $DT_{m,k}$, we have $N_{\ell_{n,k},k} - \ell_{n,k}\hat{\rho}_{\ell_{n,k},k} = \min_{1\leq j<k}\left\{N_{\ell_{n,k},j} - \ell_{n,k}\hat{\rho}_{\ell_{n,k},j}\right\}$
, then

$$N_{\ell_{n,k},k} - \ell_{n,k}\hat{\rho}_{\ell_{n,k},k} \leq 0 \tag{5.40}$$

So by gathering inequalities (5.39) and (5.40) we obtain

$$N_{\ell_{n,k},k} - \ell_{n,k}\hat{\rho}_{\ell_{n,k},k} \leq \left(\sqrt{\ell_{n,k}} - \frac{K}{2}\right)^{+} \leq \left(\sqrt{n} - \frac{K}{2}\right)^{+}$$

and by applying this result in (5.38) we get

$$\left(\frac{N_{n,k}}{n} - \hat{\rho}_{n,k}\right)^{+} \leq \left(\frac{U_{n,k} - U_{\ell_{n,k},k}}{n} + \left(\frac{1}{\sqrt{n}} - \frac{K}{2n}\right)^{+}\right) \to_{n\to+\infty} 0$$

Following the same proof steps in Lemma 5.2 we can prove that $\lim_{n\to+\infty} \frac{N_{n,k}}{n} = v$ for all $k \in [K]$. $\square$

6 Test Statistics

6.1 Pearson Chi-Square Test

Consider the homogeneity test ...

The Pearson Chi-Square statistic can be written as

$$X = \sum_{k=1}^{K} \frac{\left(S_{n,k} - N_{n,k}\hat{p}_n\right)^2}{N_{n,k}\hat{p}_n(1-\hat{p}_n)} \tag{6.1}$$

where for $j \in [K]$ we have

$$\omega_{n,j} = \frac{N_{n,j}}{n} \quad \text{and} \quad S_{n,k} = \hat{p}_{n,k}N_{n,k}, \quad \hat{p}_n = \frac{\sum_{k=1}^{K} S_{n,k}}{n} = \sum_{k=1}^{K}\omega_{n,k}\hat{p}_{n,k}$$

So we rewrite the test statistic as

$$\begin{aligned} X &= \sum_{k=1}^{K} \frac{N_{n,k}\left(\hat{p}_{n,k} - \hat{p}_n\right)^2}{\hat{p}_n(1-\hat{p}_n)} \\ &= \frac{n}{\hat{p}_n(1-\hat{p}_n)} \sum_{k=1}^{K} \omega_{n,k}\left(\hat{p}_{n,k} - \hat{p}_n\right)^2 \end{aligned}$$

Let the following notations

- $Y_n = (\sqrt{n}\,(\hat{p}_{n,1} - p), \ldots, \sqrt{n}\,(\hat{p}_{n,K} - p))$
- $s_n = \left(\sqrt{\omega_{n,1}}, \cdots, \sqrt{\omega_{n,K}}\right), \quad P_n = s_n s_n^T, \quad \|s_n\| = 1$
- $D_n = \operatorname{diag}\left(\sqrt{\omega_{n,1}}, \ldots, \sqrt{\omega_{n,K}}\right)$
- $Z_n = \left(\sqrt{n\,\omega_{n,1}}\,(\hat{p}_{n,1} - \hat{p}_n), \ldots, \sqrt{n\,\omega_{n,K}}\,(\hat{p}_{n,K} - \hat{p}_n)\right)$

For $k \in [K]$ we have:

$$\begin{aligned}[Z_n]_k &= \sqrt{n\,\omega_{n,k}}\,(\hat{p}_{n,k} - \hat{p}_n) \\ &= \sqrt{\omega_{n,k}}\,\sqrt{n}(\hat{p}_{n,k} - p) - \sqrt{\omega_{n,k}}\,\sqrt{n}(\hat{p}_n - p)\end{aligned} \tag{6.2}$$

And we have:

$$\begin{aligned}\sqrt{n}(\hat{p}_n - p) &= \sum_{j=1}^{K} \omega_{n,j}\,\sqrt{n}(\hat{p}_{n,j} - p) \\ &= s_n^T D_n Y_n\end{aligned}$$

which yields $\sqrt{\omega_{n,k}}\,\sqrt{n}(\hat{p}_n - p) = \left[s_n \cdot s_n^T D_n Y_n\right]_k$, besides it is obvious that $\sqrt{\omega_{n,k}}\,\sqrt{n}(\hat{p}_{n,k} - p) = [D_n Y_n]_k$, replacing in (6.2) gives

$$\begin{aligned}Z_n &= D_n Y_n - s_n s_n^T D_n Y_n \\ &= (I - s_n s_n^T) D_n Y_n \\ &= (I - P_n) D_n Y_n\end{aligned}$$

So:

$$\begin{aligned}X &= \frac{1}{\hat{p}_n(1 - \hat{p}_n)}\,\|Z_n\|^2 \\ &= \frac{p(1-p)}{\hat{p}_n(1 - \hat{p}_n)} \left\|(I - P_n)\,\frac{D_n Y_n}{\sqrt{p(1-p)}}\right\|^2 \\ &= \frac{p(1-p)}{\hat{p}_n(1 - \hat{p}_n)} \frac{(D_n Y_n)^T}{\sqrt{p(1-p)}}(I - P_n)\frac{D_n Y_n}{\sqrt{p(1-p)}}\end{aligned}$$

Also, we have that $\omega_{n,j} = \frac{N_{n,j}}{n} \to v_j$, then

$$\frac{D_n}{\sqrt{p(1-p)}} \xrightarrow{p} \frac{1}{\sqrt{p(1-p)}} D, \quad D := \operatorname{diag}(\sqrt{v_1}, \ldots, \sqrt{v_K}) \tag{6.3}$$

And by assumption:

$$Y_n \xrightarrow{d} \sqrt{p(1-p)}\,D^{-1} G, \quad \text{where } G \sim \mathcal{N}(0, I_K) \tag{6.4}$$

Using Slutsky on (6.3) and (6.4) we obtain that

$$\left(\frac{D_n}{\sqrt{p(1-p)}}, Y_n\right) \xrightarrow{d} \left(\frac{D}{\sqrt{p(1-p)}}, \sqrt{p(1-p)} D^{-1} G\right) = (D, D^{-1} G)$$

So using Continuous Mapping Theorem on the map $(y, a) \mapsto \frac{1}{\sqrt{p(1-p)}} a y$, we get

$$\frac{D_n Y_n}{\sqrt{p(1-p)}} \xrightarrow{d} G \sim \mathcal{N}(0, I_K) \tag{6.5}$$

Now using that $s_n \to s := (\sqrt{v_1}, \dots, \sqrt{v_K})$ we get

$$I - P_n \overset{p}{\to} I - ss^T \tag{6.6}$$

Again using Slutsky on (6.5) and (6.6), then using continuous mapping theorem on the map $(y, a) \mapsto y^T ay$, we get

$$\frac{(D_n Y_n)^T}{\sqrt{p(1-p)}}(I - P_n)\frac{D_n Y_n}{\sqrt{p(1-p)}} \xrightarrow{d} G^T(I - ss^T)G \tag{6.7}$$

and finally as $\frac{p(1-p)}{\hat{p}_n(1-\hat{p}_n)} \xrightarrow{p} 1$ we use Slutsky again with (6.7) and that $I - ss^T$ is an orthogonal projector with rank $K-1$, we conclude that

$$X_n \xrightarrow{d} G^T(I - ss^T)G \ \sim \ \chi^2_{K-1}$$

7 Conclusion

In this work we revisited the theory of response-adaptive randomization, focusing on the gap between the general flexibility of the DBCD and the efficiency of the **ERADE**. Our main contribution was to extend the **ERADE** framework beyond the two-treatment case and to propose a general algorithm that accommodates K treatments while preserving the desirable asymptotic properties of convergence, normality, and efficiency. This provides, to the best of our knowledge, the first construction of an **ERADE**-type design for multiple arms.

Beyond clinical trials, we highlighted the natural connections between adaptive randomization and the multi-armed bandit literature, particularly the problem of best-arm identification. The proposed design fits within this framework and opens new perspectives for analyzing adaptive procedures in terms of sample complexity, stopping rules, and optimality criteria that are central in sequential learning.

Recent independent work has confirmed the importance of this direction by extending **ERADE** to multiple treatments. Our results complement these developments by providing a broader and more general framework, and by explicitly linking the analysis to the multi-armed bandit setting.

Several questions remain open. In particular, the precise implications of our asymptotic normality results for stopping rule design and finite-sample performance deserve further study. Overall, this work contributes to bridging two active fields: the statistical design of clinical trials and the theory of sequential decision making.

References

Chow, Y. S. (1967). On a strong law of large numbers for martingales. *Ann. Math. Statist.*, 328:610–611.

Doob, J. L. (1936). Note on probability. *Ann. Math*, 37:363–367.

Eisele, J. (1994). The doubly adaptive biased coin design for sequential clinical trials. *Journal of Statistical Planning and Inference*, 38:249–262.

Eisele, J. and Woodroofe, M. (1995). Central limit theorems for doubly adaptive biased coin designs. *Annals of Statistics*, 23:234–254.

Gall, J.-F. L. (2006). Intégration, probabilités et processus aléatoires. Cours de Master 2, Université Paris-Sud. Lecture notes.

Garivier, A. and Kaufmann, E. (2016). Optimal best arm identification with fixed confidence. In *Proceedings of the 29th Conference On Learning Theory*.

Hall, P. and Heyde, C. C. (1980). Martingale limit theory and its applications. *Academic Press, London*.

Hu, F. and Rosenberger, W. F. (2003). Optimality, variability, power: Evaluating response adaptive randomization procedures for treatment comparisons. *Journal of the American Statistical Association*.

Hu, F. and Zhang, L.-X. (2004). Asymptotic proerties of doubly adaptive biased coin designs for multitreatment clinical trials. *Annals of Statistics*, 32(1).

Hu, F., Zhang, L.-X., and He, X. (2006). The theory of response-adaptive randomization in clinical trials. *Wiley, New York*.

Hu, F., Zhang, L.-X., and He, X. (2009). Efficient randomized-adaptive designs. *Annals of Statistics*, 37(5A):2543–2560.

Kallenberg, O. (2021). *Foundations of Modern Probability*, volume 275 of *Probability Theory and Stochastic Modelling*. Springer, New York, 3rd edition.

Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion*, volume 293 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, 3rd edition.

Stout, W. F. (1970). A martingale analogue of kolmogorov's law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 15:279–290.

A DBCD Adaptive Design

Lemma A.1 (cf. Lemma A.1 in Hu and Zhang (2004)). *Let $B_{n,n} = I$ and $B_{n,i} = \prod_{j=i}^{n-1}\left(I + j^{-1}H\right)$. If two sequences of matrices (Q_n) and (P_n) satisfy*

$$Q_n \;=\; P_n + \sum_{k=1}^{n-1} \frac{Q_k}{k}\, H,$$

that is,

$$Q_n \;=\; \Delta P_n + Q_{n-1}\left(I + \frac{H}{n-1}\right), \qquad n \geq 2,$$

where $\Delta P_1 = P_1$ and $\Delta P_n = P_n - P_{n-1}$ for $n \geq 2$, then

$$Q_n = \sum_{m=2}^{n} \Delta P_m B_{n,m} + Q_1 B_{n,1} \;=\; (Q_1 - P_1)B_{n,1} + \sum_{m=1}^{n} \Delta P_m B_{n,m} \tag{A.1}$$

$$= (Q_1 - P_1)B_{n,1} + P_n + \sum_{m=1}^{n-1} P_m \frac{H}{m}\, B_{n,m+1}. \tag{A.1}$$

Also,

$$\|B_{n,m}\| \;\leq\; C\,(n/m)^{\lambda}\, \log^{\nu-1}(n/m) \quad \textit{for all } m = 1, \ldots, n,\ n \geq 1, \tag{A.2}$$

where, in the rest of the paper, $\log x = \ln(x \vee e)$.

Lemma A.2 (cf. Lemma A.2 in Hu and Zhang (2004)). *If two sequences of matrices (Q_n) and (P_n) satisfy*

$$\Delta Q_n \;=\; \Delta P_n + \frac{Q_{n-1}}{n-1}\, H + o\left(\frac{\|Q_{n-1}\|}{n-1}\right),$$

then, for any $\beta > 0$,

$$Q_n \;=\; O(1) \sum_{m=1}^{n} \|P_m\|\, \frac{1}{m} \left(\frac{1}{m}\right)^{\lambda+\beta}.$$

Lemma A.3 (From Lyapunov to Lindeberg Condition). *Let $X_1, X_2, \ldots$ be independent random variables with $\mathbb{E}[X_k] = 0$ and finite variances $\sigma_k^2 = \mathbb{E}\left[X_k^2\right]$. Set $s_n^2 := \sum_{k=1}^{n} \sigma_k^2$*

Suppose that for some $\varepsilon > 0$ the Lyapunov condition holds:

$$\frac{1}{s_n^{2+\varepsilon}} \sum_{k=1}^{n} \mathbb{E}\left[|X_k|^{2+\varepsilon}\right] \xrightarrow[n\to\infty]{} 0.$$

Then the Lindeberg condition holds:

$$\forall \eta > 0: \ \frac{1}{s_n^2} \sum_{k=1}^{n} \mathbb{E}\left[X_k^2\, \mathbf{1}\{|X_k| > \eta s_n\}\right] \xrightarrow[n\to\infty]{} 0.$$

Proof. Fix $\eta > 0$. For any real x and $\varepsilon > 0$,

$$x^2 \mathbf{1}\{|x| > \eta s_n\} \;=\; x^2 \mathbf{1}\{|x|^{\varepsilon} > (\eta s_n)^{\varepsilon}\} \;\leq\; \frac{|x|^{2+\varepsilon}}{(\eta s_n)^{\varepsilon}}$$

Taking expectations of the previous inequality applied to X_k and summing over k gives

$$\sum_{k=1}^{n} \mathbb{E}\left[X_k^2 \mathbf{1}\{|X_k| > \eta s_n\}\right] \;\leq\; \frac{1}{(\eta s_n)^{\varepsilon}} \sum_{k=1}^{n} \mathbb{E}|X_k|^{2+\varepsilon}$$

$$\Longrightarrow \ \frac{1}{s_n^2} \sum_{k=1}^{n} \mathbb{E}\left[X_k^2 \mathbf{1}\{|X_k| > \eta s_n\}\right] \;\leq\; \frac{1}{\eta^{\varepsilon}} \frac{1}{s_n^{2+\varepsilon}} \sum_{k=1}^{n} \mathbb{E}|X_k|^{2+\varepsilon} \xrightarrow[n\to\infty]{} 0,$$

by the Lyapunov assumption which gives precisely Lindeberg's condition. □

Lemma A.4. *Under the conditions of Theorem 3.3 we have for* $\delta > 0$

$$o\Big(\|\hat{\Theta}_n - \Theta\|^{1+\delta}\Big) = o\Big(n^{-\frac{1}{2}-\frac{\delta}{3}}\Big) \quad a.s. \tag{A.2}$$

$$o\left(\left\|\tfrac{N_{n-1}}{n} - v\right\|^{1+\delta}\right) = o\left(n^{-\frac{1}{2}-\frac{\delta}{3}}\right) \quad a.s. \tag{A.3}$$

Proof. (i) We have for some $\varepsilon_n \to 0$

$$\begin{aligned} o\Big(\|\hat{\Theta}_n - \Theta\|^{1+\delta}\Big) &= \varepsilon_n\, n^{-\frac{1+\delta}{2}} (\log\log n)^{\frac{1+\delta}{2}} \\ &= n^{-\frac{1}{2}-\frac{\delta}{3}} \underbrace{\varepsilon_n\, n^{-\delta/6} (\log\log n)^{\frac{1+\delta}{2}}}_{\to 0} \\ &= o\Big(n^{-\frac{1}{2}-\delta/3}\Big) \quad a.s. \end{aligned}$$

(ii) We have for some $\varepsilon_n \to 0$

$$\begin{aligned} o\left(\left\|\tfrac{N_{n-1}}{n} - v\right\|^{1+\delta}\right) &= o\left(\frac{\|N_{n-1} - nv\|^{1+\delta}}{n^{1+\delta}}\right) \\ &= o\left(\frac{(n\log\log n)^{\frac{1+\delta}{2}}}{n^{1+\delta}}\right) \\ &= \varepsilon_n\, n^{-\frac{1+\delta}{2}} (\log\log n)^{\frac{1+\delta}{2}} \\ &= n^{-\frac{1}{2}-\frac{\delta}{3}} \cdot \varepsilon_n\, \underbrace{n^{-\delta/6}(\log\log n)^{\frac{1+\delta}{2}}}_{\to 0} \\ &= o\left(n^{-\frac{1}{2}-\frac{\delta}{3}}\right) \quad \text{a.s.} \end{aligned}$$

□

Proof of Theorem 3.3. Let $Q_n = \sum_{m=1}^n \Delta Q_m$ where $\Delta Q_m = \big(\Delta Q_{m,1}, \dots, \Delta Q_{m,K}\big)$ and for all $k \in [K]$

$$Q_{m,k} = X_{m,k} \frac{\xi_{m,k} - \theta_k}{v_k}.$$

$\{Q_n\}_{n\geq 1}$ is a sequence of martingales, indeed

$$\begin{aligned} \mathbb{E}[Q_{n,k} \mid \mathcal{F}_{n-1}] &= Q_{n-1,k} + \mathbb{E}[\Delta Q_{n,k} \mid \mathcal{F}_{n-1}] \\ &= Q_{n-1,k} + \mathbb{E}\left[X_{n,k} \frac{\xi_{n,k} - \theta_k}{v_k} \middle| \mathcal{F}_{n-1}\right] \\ &= Q_{n-1,k} + \frac{1}{v_k}\, \mathbb{E}[X_{n,k} \mid \mathcal{F}_{n-1}]\, \underbrace{\mathbb{E}[\xi_{n,k} - \theta_k \mid \mathcal{F}_{n-1}]}_{=0} \\ &= Q_{n-1,k} \end{aligned}$$

So by LIL for martingales we have

$$Q_n = O\Big(\sqrt{n\log\log n}\Big) \quad \text{a.s.} \tag{A.4}$$

and we have from (3.17)

$$\begin{aligned}
\hat{\theta}_{n,k} - \theta_k &= \frac{1}{N_{n,k}+1}\left\{\sum_{m=1}^{n} X_{m,k}\,\xi_{m,k} + \theta_{0,k}\right\} - \theta_k \\
&= \frac{v_k}{N_{n,k}+1}\left\{\sum_{m=1}^{n} X_{m,k}\frac{\xi_{m,k}-\theta_k}{v_k^2}\right\} + \frac{\theta_{0,k}-\theta_k}{N_{n,k}+1} \\
&= \frac{v_k}{N_{n,k}+1}.Q_{n,k} + \frac{\theta_{0,k}-\theta_k}{N_{n,k}+1} \\
&= \frac{Q_{n,k}}{n} + \left(\frac{v_k}{N_{n,k}+1} - \frac{1}{n}\right) Q_{n,k} + \frac{\theta_{0,k}-\theta_k}{N_{n,k}+1} \\
&= \frac{Q_{n,k}}{n} + O\left(\frac{\sqrt{n\log\log n}}{n^2}\right) + O\left(\sqrt{n\log\log n}\right) + O\left(\frac{1}{n}\right) \\
&= \frac{Q_{n,k}}{n} + O\left(\frac{\log\log n}{n}\right) \quad a.s.
\end{aligned}$$

Then

$$\hat{\Theta}_n - \Theta = \frac{Q_n}{n} + O\left(\frac{\log\log n}{n}\right) \quad a.s. \tag{A.5}$$

By condition (C2) we have

$$\begin{aligned}
\hat{\rho}_n - v &= (\hat{\theta}_n - \theta)\cdot \nabla(\rho)\big|_{\Theta} + o\left(\|\hat{\Theta}_n - \Theta\|^{1+\delta}\right) \\
&= \frac{Q_n}{n}\cdot \nabla(\rho)\big|_{\Theta} + O\left(\frac{\log\log n}{n}\right) + o\left(\|\hat{\Theta}_n - \Theta\|^{1+\delta}\right)
\end{aligned}$$

For $\delta \in (0, \frac{3}{2})$ we have that $O\left(\frac{\log\log n}{n}\right) = o\left(n^{-1/2-\delta/3}\right)$ a.s., with result of (A.2) we get

$$\hat{\rho}_n - v = \frac{Q_n}{n}\,\nabla(\rho)\big|_{\Theta} + o\left(n^{-1/2-\delta/3}\right) \quad \text{a.s.} \tag{A.6}$$

On the other hand we have

$$\begin{aligned}
\Delta\big(N_n - nv\big) &= X_n - v \\
&= \Delta M_n + g\left(\frac{N_{n-1}}{n-1}, \hat{\rho}_{n-1}\right) - g(v,v) \\
&= \Delta M_n + \left(\frac{N_{n-1}}{n-1} - v\right) H + (\hat{\Theta}_{n-1} - \Theta)\,\nabla(\rho)\big|_{\Theta}\,E + o\left(\left\|\tfrac{N_{n-1}}{n-1} - v\right\|^{1+\delta}\right) \\
&\quad + o\left(\|\hat{\Theta}_{n-1} - \Theta\|^{1+\delta}\right) \quad \text{a.s.} \\
&= \Delta M_n + \frac{Q_n}{n}\,\nabla(\rho)\big|_{\theta}\,E + \frac{N_{n-1} - (n-1)v}{n-1}\,H + o\left(n^{-\frac{1}{2}-\frac{\delta}{3}}\right) + o\left(\left\|\tfrac{N_{n-1}}{n-1} - v\right\|^{1+\delta}\right) \\
&\quad + o\left(\|\hat{\Theta}_{n-1} - \Theta\|^{1+\delta}\right) \quad \text{a.s.}
\end{aligned}$$

So using (A.3) we obtain

$$N_n - nv = \Delta M_n + \frac{Q_n}{n}\,\nabla(\rho)\big|_{\theta}\,E + (N_{n-1} - (n-1)v)\,\left(I + \frac{H}{n-1}\right) + o\left(n^{-\frac{1}{2}-\frac{\delta}{3}}\right) \quad \text{a.s.} \tag{A.7}$$

Now we can use Lemma A.1 as we have similar composition in the previous equation, we get with the notations defined in Lemma A.1

$$N_n - nv = \sum_{m=1}^{n}\left(\Delta M_m + \frac{Q_m}{m}\nabla(\rho)\big|_\theta \cdot E + o\Big(m^{-1/2-\delta/3}\Big)\right) B_{n,m}$$
$$= \sum_{m=1}^{n}\Delta M_m B_{n,m} + \sum_{m=1}^{n}\sum_{k=1}^{m}\frac{\Delta Q_k}{m}\nabla(\rho)\big|_\Theta \cdot E\, B_{n,m} + \sum_{m=1}^{n} o\Big(m^{-1/2-\delta/3}\Big)\left(\frac{n}{m}\right)^\lambda \log^{\nu-1}\left(\frac{n}{m}\right)$$
$$= \sum_{m=1}^{n}\Delta M_m B_{n,m} + \sum_{m=1}^{n}\Delta Q_m \nabla(\rho)\big|_\theta \cdot E \sum_{k=m}^{n}\frac{1}{k}B_{n,k} + o\Big(n^{-1/2-\delta/3}\Big) \quad \text{a.s.}$$

Since

$$\sum_{m=1}^{n} o\Big(m^{-1/2-\delta/3}\Big)\left(\frac{n}{m}\right)^\lambda \log^{\nu-1}\left(\frac{n}{m}\right) = o\Big(n^{-1/2-\delta/3}\Big), \quad \text{a.s.}$$

so we obtain for $U_n := \sum_{m=1}^{n}\Delta M_m B_{n,m} + \sum_{m=1}^{n}\Delta Q_m \nabla(\rho)\big|_\theta \cdot E \sum_{k=m}^{n}\frac{1}{k}B_{n,k}$.

$$N_n - nv = U_n + o\Big(n^{-1/2-\delta/3}\Big) \quad \text{a.s.} \tag{A.8}$$

Note. The presence of the term $o\big(n^{-1/2-\delta/3}\big)$ in (A.6) and (A.8) is for having a vanishing term when multiplying by $\sqrt{n}$ when using CLT as presented in Hall and Heyde (1980).

Now we check Lindberg condition for U_n based on Lemma A.3, by condition (A2) and Lemma A.1 result we have for $0 < \varepsilon < \frac{1}{\lambda} - 2$:

$$\frac{1}{n^{1+\varepsilon/2}}\left\{\left\|\Delta M_n\, B_{n,m}\right\|^{2+\varepsilon} + \|\Delta Q_m\|^{2+\varepsilon} + \left\|\Delta Q_m \nabla(\rho)\big|_\theta \cdot E \sum_{k=m}^{n}\frac{1}{k}B_{n,k}\right\|^{2+\varepsilon}\right\}$$
$$\le \frac{C}{n^{1+\varepsilon/2}}\sum_{m=1}^{n}\left\{\left(\left(\frac{n}{m}\right)^\lambda \log^{1-\nu}\left(\tfrac{n}{m}\right)\right)^{2+\varepsilon} + \left(\sum_{k=m}^{n}\frac{1}{k}\left(\tfrac{n}{k}\right)^\lambda \log^{\nu-1}\left(\tfrac{n}{k}\right)\right)^{2+\varepsilon}\right\}$$
$$\le \frac{C}{n^{1+\varepsilon/2}}\sum_{m=1}^{n}\left(\left(\frac{n}{m}\right)^\lambda \log^{1-\nu}\left(\tfrac{n}{m}\right)\right)^{2+\varepsilon}$$
$$\le \frac{C}{n^{\varepsilon/2}} \longrightarrow 0, \quad n \to +\infty$$

With Lindberg condition being verified, we can use Hall and Heyde (1980) applied on (A.6) and (A.8) to conclude. For the calculation of covariance matrices the reader can refer to Hu and Zhang (2004). ☐

B Generalized ERADE for K treatments

Lemma B.1. *Let $\alpha \in [0,1)$, $v \in (0,1)$, $C > 0$, $(\ell_n)_{n\ge1}$ a positive increasing sequence such that $\forall n \in \mathbb{N} : \ell_n \le n$ and (U_n) a real sequence such that $U_n \sim -n\alpha v$, then:*

$$\lim_{n\to+\infty}\left(\frac{C}{n} + \frac{U_n - U_{\ell_n}}{n}\right)^+ \longrightarrow 0.$$

Proof. **Case 1:** $(\ell_n)_{n\in\mathbb{N}}$ is bounded

As (ℓ_n) is increasing sequence then ℓ_n is stationary so $\frac{U_{\ell_n}}{n} \longrightarrow 0$, which gives:

$$\lim_{n\to+\infty} \frac{C}{n} + \frac{U_n - U_{\ell_n}}{n} = -\alpha v \implies \lim_{n\to+\infty} \left(\frac{C}{n} + \frac{U_n - U_{\ell_n}}{n}\right)^+ = 0$$

Case 2: $\ell_n \to +\infty$

In this case we have $\lim_{n\to+\infty} \frac{U_{\ell_n}}{\ell_n} = -\alpha v$, then:

$$\begin{aligned}
\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geq N, \quad -\varepsilon \leq \frac{U_{\ell_n}}{\ell_n} + \alpha v \leq \varepsilon \quad \text{a.s.} \\
\Rightarrow -\varepsilon - \alpha v \leq \frac{U_{\ell_n}}{\ell_n} \leq \varepsilon - \alpha v \quad \text{a.s.} \\
\Rightarrow -\frac{U_{\ell_n}}{n} \leq \varepsilon\frac{\ell_n}{n} + \alpha v\frac{\ell_n}{n} \leq \varepsilon + \alpha v \quad \text{a.s.} \\
\Rightarrow \frac{C}{n} + \frac{U_n - U_{\ell_n}}{n} \leq \frac{C}{n} + \frac{U_n}{n} + \varepsilon + \alpha v \quad \text{a.s.}
\end{aligned}$$

So as $U_n \sim -n\alpha v$ we obtain that $\lim_{n\to+\infty} \frac{C}{n} + \frac{U_n - U_{\ell_n}}{n} \leq 0$ then $\lim_{n\to+\infty} \left(\frac{C}{n} + \frac{U_n - U_{\ell_n}}{n}\right)^+ = 0$

□

Lemma B.2. *Suppose $\{M_n, \mathcal{F}_n; n \geq 1\}$ is a martingale with $\mathbb{E}|\Delta M_n|^2 \leq C_0$, where $\Delta M_n = M_n - M_{n-1}$ and $\{\mathcal{F}_n\}$ is a filter of sigma fields with $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$. Then, for any $L > 4$,*

(i)

$$\mathbb{E}\left[\max_{m\leq L} |M_n - M_{n-m}|\right] \leq 4\sqrt{C_0 L} \tag{B.1}$$

(ii)

$$\mathbb{E}\left[\max_{m: L\leq m\leq n} \frac{|M_n - M_{n-m}|}{m}\right] \leq 12\sqrt{\frac{C_0}{L}} \tag{B.2}$$

Proof. (i) The triangular inequality gives that $|M_n - M_{n-m}| \leq |M_n - M_{n-L}| + |M_{n-m} - M_{n-L}|$, then

$$\begin{aligned}
\sup_{1\leq m\leq L} |M_n - M_{n-m}| &\leq |M_n - M_{n-L}| + \sup_{1\leq m\leq L} |M_{n-m} - M_{n-L}| \\
&\leq 2 \sup_{k\in\{n-L+1,\ldots,n\}} |M_k - M_{n-L}| \\
&= 2 \sup_{k\in\{1,\ldots,L\}} |S_k|
\end{aligned} \tag{B.3}$$

where $S_k := M_{n-L+k} - M_{n-L}$.

Define now the new forward filtration $\mathcal{G}_k := \mathcal{F}_{n-L+k}$. We have

$$\begin{aligned}
\mathbb{E}[S_k \mid \mathcal{G}_{k-1}] &= S_{k-1} + \mathbb{E}[\Delta S_k \mid \mathcal{G}_{k-1}] \\
&= S_{k-1} + \mathbb{E}[M_{n-L+k} - M_{n-L+k-1} \mid \mathcal{G}_{k-1}] \\
&= S_{k-1}
\end{aligned}$$

so $\{S_k\}$ is a $\mathcal{G}_k$-martingale. Then $\{S_k^2\}$ is a sub-martingale. Applying Doob's maximal inequality as in Lemma C.3 we obtain

$$\begin{aligned}
\mathbb{P}\left(\sup_{k\in\{1,\ldots,L\}} S_k^2 \geq x^2\right) &\leq \frac{\mathbb{E}[S_L^2]}{x^2} \\
\implies \mathbb{P}\left(\sup_{k\in\{1,\ldots,L\}} |S_k| \geq x\right) &\leq \frac{\mathbb{E}[S_L^2]}{x^2}
\end{aligned} \tag{B.4}$$

$\{S_k\}$ being a martingale gives that $\{\Delta S_k\}$ are orthogonal, i.e. $\mathrm{Cov}(\Delta S_k, \Delta S_j) = 0$. So:

$$\begin{aligned}
\mathbb{E}[S_L^2] &= \sum_{k=1}^{L} \mathbb{E}[\,|\Delta S_k|^2\,] \\
&= \sum_{k=1}^{L} \mathbb{E}\big[\,|\Delta M_{n-L+k}|^2\,\big] \\
&\leq LC_0.
\end{aligned}$$

So from (B.3) and (B.4) we obtain

$$\mathbb{P}\left(\sup_{m\in\{1,\dots,L\}} |M_n - M_{n-m}| \geq x \right) \leq 4\frac{LC_0}{x^2}$$

Then

$$\begin{aligned}
\mathbb{E}\left[\sup_{m\in\{1,\dots,L\}} |M_n - M_{n-m}| \right] &= \int_0^{+\infty} \mathbb{P}\left(\sup_{m\in\{1,\dots,L\}} |M_n - M_{n-m}| \geq x \right) dx \\
&\leq \int_0^{2\sqrt{LC_0}} 1\dot{d}x + 4\int_{2\sqrt{LC_0}}^{+\infty} \frac{LC_0}{x^2}\dot{d}x \\
&= 4\sqrt{LC_0}
\end{aligned}$$

(ii)

$$\begin{aligned}
\mathbb{E}\left[\max_{L\leq m\leq n} \frac{|M_n - M_{n-m}|}{m} \right] &\leq \sum_{\log_2(L)\leq j\leq \log_2(n)} \frac{1}{2^{j-1}} \sqrt{2}\,\mathbb{E}\left[\max_{2^{j-1}\leq m<2^j} |M_n - M_{n-m}| \right] \\
&\leq \sum_{j\geq \log_2(L)} \frac{1}{2^{j-1}}\,\mathbb{E}\left[\max_{m<2^j} |M_n - M_{n-m}| \right] \\
&\leq \sum_{j\geq \log_2(L)} \frac{4\sqrt{C_0\cdot 2^j}}{2^{j-1}} \\
&= 4\sum_{j\geq \log_2(L)} \frac{\sqrt{C_0}}{\sqrt{2^{j-2}}} \leq 12\sqrt{\frac{C_0}{L}}
\end{aligned}$$

□

Lemma B.3. *For all $\ell \geq 1, L < n$ and any sequence $(Q_n) \subset \mathbb{R}^d$, we have*

$$\|Q_n - Q_\ell\| \leq (n-\ell) \max_{L\leq m\leq n} \frac{\|Q_n - Q_{n-m}\|}{m} + \max_{m<L} \|Q_n - Q_{n-m}\|.$$

Proof. Let $d = n - \ell$. We consider two cases.

If $d \leq L$:

$$\|Q_n - Q_\ell\| = \|Q_n - Q_{n-d}\| \leq \max_{m<L} \|Q_n - Q_{n-m}\|.$$

If $d > L$:

$$\begin{aligned}
\frac{\|Q_n - Q_\ell\|}{d} &= \frac{\|Q_n - Q_{n-d}\|}{d} \\
&\leq \max_{L\leq m\leq n} \frac{\|Q_n - Q_{n-m}\|}{m}
\end{aligned}$$

$$\implies \|Q_n - Q_\ell\| \leq (n-\ell) \max_{L\leq m\leq n} \frac{\|Q_n - Q_{n-m}\|}{m}$$

So gathering the two upper bounds for both cases we get:

$$\|Q_n - Q_\ell\| \leq (n-\ell) \max_{L\leq m\leq n} \frac{\|Q_n - Q_{n-m}\|}{m} + \max_{m<L} \|Q_n - Q_{n-m}\|,$$

$\square$

C General Probability and Martingales Results

Theorem C.1 (SLLN for Martingales; cf. Chow (1967)). *Let $(X_n, \mathcal{F}_n)_{n\geq 1}$ be a martingale difference sequence, i.e.*

$$\mathbb{E}[X_n \mid \mathcal{F}_{n-1}] = 0, \quad n \geq 1.$$

Suppose there exists $\delta > 0$ such that

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\big(|X_n|^{2+\delta}\big)}{n^{1+\delta}} < \infty.$$

Then, with $M_n = \sum_{k=1}^{n} X_k$, we have

$$\frac{M_n}{n} \longrightarrow 0 \quad a.s.$$

Equivalently,

$$M_n = o(n) \quad a.s.$$

Theorem C.2 (LIL for Martingale; cf. Stout (1970)). *Let $(M_n, \mathcal{F}_n)_{n\geq 0}$ be a square-integrable martingale with martingale differences $Y_n = M_n - M_{n-1}$, $M_0 = 0$. Define the predictable quadratic variation*

$$s_n^2 = \sum_{k=1}^{n} \mathbb{E}\left[Y_k^2 \,\middle|\, \mathcal{F}_{k-1}\right],$$

and set

$$u_n = \sqrt{2\log\log s_n^2}.$$

Assume that $s_n^2 \to \infty$ almost surely and that the increments satisfy a Lindeberg-type condition (for example, $\max_{1\leq k\leq n} |Y_k| = o(s_n/u_n)$ a.s.). Then

$$\limsup_{n\to\infty} \frac{M_n}{s_n u_n} = 1 \quad \text{and} \quad \liminf_{n\to\infty} \frac{M_n}{s_n u_n} = -1, \quad a.s.$$

In particular, if $s_n^2 \sim \sigma^2 n$ for some $\sigma^2 > 0$, then

$$M_n = O\big(\sqrt{n \log\log n}\big) \quad a.s.$$

Lemma C.1 (Borel-Cantelli Lemma; cf. Corollary 9.21 in Kallenberg (2021)). *For a filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ and $A_n \in \mathcal{F}_n$ for $n \in \mathbb{N}$ we have*

$$\{A_n \quad i.o.\} = \left\{ \sum_{n\geq 1} \mathbb{P}\left(A_n \mid \mathcal{F}_{n-1}\right) = +\infty \right\} \quad a.s.$$

Lemma C.2 (Doob Maximal Inequality; cf. Theorem 12.4.2 in Gall (2006)). *Soit $\{X_n\}_{n\in\mathbb{N}}$ a submartingale. Then for all $c > 0$ and $n \in \mathbb{N}$,*

$$\mathbb{P}\left(\sup_{0\leq k\leq n} X_k \geq c\right) \leq \frac{\mathbb{E}\left[X_n^+\right]}{c}$$

Lemma C.3 (Doob Maximal Inequality; cf. Corollary 2.1.6 in Revuz and Yor (1999))**.** *Let $\{X_n\}_{n\geq 1}$ a martingale or a positive sub-martingale. Then for every $p \geq 1$ and $c > 0$,*

$$\mathbb{P}\left(\sup_{1\leq k\leq n} |X_k| \geq \lambda\right) \leq \frac{\mathbb{E}\left[|X_n|^p\right]}{\lambda^p}$$

and for any $p > 1$

$$\mathbb{E}\left[\sup_{1\leq k\leq n} |X_k|^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[|X_n|^p\right]$$

In particular for $p = 2$ we have,

$$\mathbb{E}\left[\sup_{1\leq k\leq n} X_k^2\right] \leq 4\,\mathbb{E}\left[X_n^2\right]$$

Lemma C.4. *Suppose $\{X_n\}_{n\geq 1}$ is a square-integrable martingale with $\mathbb{E}[X_1] = 0$. Then for $\lambda > 0$:*

$$P\left(\sup_{1\leq i\leq n} X_i \geq \lambda\right) \leq \frac{\mathrm{Var}(X_n)}{\mathrm{Var}(X_n) + \lambda^2}.$$

Proof. For $\alpha > 0$ we have

$$\mathbb{P}\left(\sup_{0\leq k\leq n} X_k \geq c\right) = \mathbb{P}\left(\sup_{0\leq k\leq n} (X_k + \alpha)^2 \geq (c+\alpha)^2\right)$$

As $\{X_n\}$ is a martingale then $\{X_n^2\}$ is a sub-martingale, so we can apply Lemma C.2 giving that

$$\begin{aligned}
\mathbb{P}\left(\sup_{0\leq k\leq n} (X_k + \alpha)^2 \geq (c+\alpha)\right) &\leq \frac{\mathbb{E}\left[(X_n + c)^2\right]}{(c+\alpha)^2} \\
&= \frac{\mathbb{V}ar\left[X_n\right] + \alpha^2}{(c+\alpha)^2} \\
&= \frac{\mathbb{V}ar\left[X_n\right] + \left(\frac{\mathbb{V}ar[X_n]}{c}\right)^2}{\left(c + \frac{\mathbb{V}ar[X_n]}{c}\right)^2}; \text{ Taking } \alpha = \frac{\mathbb{V}ar\left[X_n\right]}{c} \\
&= \frac{\mathbb{V}ar\left[X_n\right]}{\mathbb{V}ar\left[X_n\right] + c^2}
\end{aligned}$$

which proves the claim.

∎

Lemma C.5. *Let $(u_n)_{n\geq 1}$ be a sequence of positive numbers and let $\{X_n\}_{n\geq 1}$ be real-valued random variables with*

$$\mathbb{E}|X_n| < \infty \quad \text{and} \quad \mathbb{E}|X_n| \leq \frac{C}{u_n} \qquad \text{for all } n,$$

for some constant $C < \infty$. Then

$$X_n = O_p(u_n),$$

i.e., for every $\varepsilon > 0$ there exists $M < \infty$ such that

$$\sup_{n \geq 1} \mathbb{P}\big(|X_n| > M u_n\big) \leq \varepsilon.$$

Proof. By Markov's inequality applied to the non-negative random variable $|X_n|$,

$$\mathbb{P}\big(|X_n| > M u_n\big) \ \leq \ \frac{\mathbb{E}|X_n|}{M u_n} \ \leq \ \frac{C}{M}.$$

Given $\varepsilon > 0$, choose $M \geq C/\varepsilon$. Then, for all n,

$$\mathbb{P}\big(|X_n| > M u_n\big) \leq \varepsilon.$$

This is precisely the definition of $X_n = O_p(u_n)$. $\square$