Dependent Randomized Rounding for Budget Constrained Experimental Design

Khurram Yamin[1] **Edward Kennedy**[1] **Bryan Wilder**[1]

[1]Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

Abstract

Policymakers in resource-constrained settings require experimental designs that satisfy strict budget limits while ensuring precise estimation of treatment effects. We propose a framework that applies a dependent randomized rounding procedure to convert assignment probabilities into binary treatment decisions. Our proposed solution preserves the marginal treatment probabilities while inducing negative correlations among assignments, leading to improved estimator precision through variance reduction. We establish theoretical guarantees for the inverse propensity weighted and general linear estimators, and demonstrate through empirical studies that our approach yields efficient and accurate inference under fixed budget constraints.

1 INTRODUCTION

Experimental design in many fields requires balancing two competing sets of objectives. On one side, practical constraints such as limited budgets or fixed resource availability demand that the total number of treated units be strictly bounded. On the other side, the design must preserve the individual-level assignment probabilities that have been determined—often based on factors such as risk or potential benefit—to ensure that the estimation of treatment effects is both statistically efficient and unbiased. In many settings, assignment probabilities for an experiment may be chosen to reflect concerns like strategy-proofness or allocating treatment to higher-need individuals. [Kido, 2023, Chassang et al., 2012, Narita, 2021, Wilder and Welle, 2024]

Consider, for example, a public health intervention during a flu outbreak where only a limited number of vaccines are available. In such a setting, independent Bernoulli randomization might preserve each individual's target probability of being treated, but it does not guarantee that the total number of vaccinations will exactly match the available supply. Conversely, methods that force a deterministic count of treatments (e.g., conditional Poisson sampling) may enforce the resource constraint, yet they risk distorting the pre-specified assignment probabilities. [Hájek, 1964, Grafström, 2005] When the intended probabilities are altered, individuals who are at higher risk and thus designed to have a greater chance of receiving the vaccine may end up being underrepresented in the treated group. Such misallocation can lead to poorer health outcomes for those who need the intervention the most. Conversely, individuals with lower risk might be overtreated, potentially wasting limited resources and causing unnecessary exposure to the treatment's side effects. In both cases, the fairness and overall effectiveness of the intervention are compromised. Moreover, from a statistical standpoint, the mis-specified weights that would be used in an inverse probability weighted (IPW) estimator no longer reflect the true design, thereby inflating the variance of effect estimates [Hájek, 1964, Grafström, 2005].

In this paper, we propose a framework based on *dependent randomized rounding*—specifically, a swap rounding algorithm as introduced in [Chekuri and Vondrák, 2009]—that addresses these challenges. Swap rounding operates by iteratively pairing fractional assignment probabilities and redistributing probability mass between the paired units in a randomized manner. This procedure continues until an integral solution is reached. Our method exactly enforces a fixed treatment count while preserving the marginal assignment probabilities determined by the experimental design. Moreover, by inducing negative correlations among the treatment indicators, our approach leads to a reduction in the variance of standard estimators, such as the inverse probability weighted (IPW) estimator. As such, we balance the competing objectives of satisfying resource constraints while maintaining an unbiased and low-variance estimator. Furthermore, our variance reduction benefits are empirically substantial enough for small samples that our technique may be of interest just as a variance reduction technique even when budget constraints are not themselves at play. The ben-

efits achieved in this regime from swap-rounding are somewhat surprising as despite the fact that the swap-rounding algorithm is solving an ostensibly similar problem, it was proposed for different purposes, primarily in the design of algorithms for submodular optimization [Chekuri and Vondrák, 2009]. To our knowledge, no previous work has made this connection between budget-constrained experimental design and combinatorial optimization. Furthermore, while the swap-rounding algorithm had been analyzed for the purposes of optimization, its statistical properties which we explore in this paper had not previously been analyzed.

The remainder of the paper is organized as follows. We describe our dependent randomized rounding mechanism as well as the causal assumptions we make. We then discuss our strategy for integrating swap rounding into IPW estimation and derive an estimator for the variance of IPW using swap-rounding. We discuss the estimator's theoretical guarantees (unbiased estimation and lower variance when IPW is used with swap rounding than without). We then detail the unbiasedness and variance reduction properties for general linear estimators under our approach. Following this, we propose a modification to swap-rounding where we prioritize swaps between units with similar covariates. To demonstrate the efficacy of our approach in empirical studies, we use a combination of synthetic and semi-synthetic examples based on real-world data in two domains: infant health and public housing. Finally, we conclude with a discussion of the broader implications for experimental design in settings with competing constraints.

2 RELATED WORKS

Experimental design under resource constraints has long been a topic of interest in both survey sampling and causal inference. Classical approaches, such as conditional Poisson sampling [Hájek, 1964], were introduced to ensure that a fixed sample size is achieved. However, these methods can distort the intended assignment probabilities and introduce bias into estimators [Aires, 2000, Grafström, 2005]. Similar challenges have been observed in unequal probability sampling [Isaki and Fuller, 1984], where the balancing of inclusion probabilities is critical to obtain reliable estimates. These limitations motivate the search for alternative methods that maintain the original probability structure while strictly enforcing allocation constraints, and that also preserve desirable properties such as unbiasedness and robustness in finite samples.

More recently, advances in combinatorial optimization have led to the development of dependent randomized rounding techniques. In particular, swap rounding [Chekuri and Vondrák, 2009] is a powerful method for converting fractional assignment solutions into integral ones while preserving marginal probabilities. This approach not only maintains unbiasedness by ensuring that the marginal assignment probabilities are exactly retained but also imposes a negative correlation structure among the rounded variables. Related work in submodular maximization [Ageev and Sviridenko, 2004] and robust approximation algorithms [Srinivasan, 2001] further demonstrates that these rounding procedures yield solutions with enhanced concentration properties, improved stability, and robustness to model misspecification, all of which are critical for reliable inference in experimental settings. Our work builds on these ideas by employing swap rounding to enforce an exact treatment count while retaining the desired statistical properties of the design, which in turn improves the precision and decreases the variance of standard estimators like the inverse propensity weighted (IPW) and doubly robust (DR) estimators.

In addition, re-randomization methods have been explored to improve covariate balance and enhance the efficiency of experimental designs [Morgan and Rubin, 2012, Li et al., 2018]. Re-randomization repeatedly randomizes the treatment assignments until a desired balance is achieved. Our framework, in contrast, directly enforces the resource constraint and connects with optimal allocation strategies such as Neyman allocation [Neyman, 1934] that similarly adjusts sampling probabilities to minimize variance. This unified approach preserves the key statistical properties of the initial design and is particularly relevant in settings like public health interventions and social programs [Athey and Imbens, 2017, Imbens and Rubin, 2015], where efficient use of limited resources can have a significant social impact.

3 PROBLEM FORMULATION

In many experimental settings, researchers are given a desired set of assignment probabilities that are designed to balance goals like welfare, efficiency, or fairness. Suppose we have a set of n candidate units. For each unit i we have a fractional assignment probability $p_i \in [0, 1]$ such that

$$\sum_{i=1}^{n} p_i = B,$$

where B represents a fixed resource constraint (for example, the exact number of vaccines available in a public health intervention). Our goal is to convert this fractional probability vector into a distribution over binary treatment assignments $A \in \{0, 1\}^n$ that meets the following criteria:

1. **Feasibility:** Every assignment A satisfies the resource constraint exactly, i.e.,

$$\sum_{i=1}^{n} A_i = B.$$

2. **Preservation of Marginals:** For each unit i, the marginal probability of treatment remains unchanged, i.e.,

$$\Pr(A_i = 1) = p_i.$$

3. **Estimation Efficiency:** The joint distribution of the binary assignment vector should be such that when used to estimate treatment effects (e.g., via an inverse probability weighted estimator), the estimator's variance is minimized compared to a baseline method (e.g., independent randomization).

In other words, the problem is to construct a randomized procedure that maps the fractional vector $p = (p_1, p_2, \ldots, p_n)$ (which exactly sums to B) to a valid binary assignment vector A that satisfies the above three criteria.

3.1 PROPOSED APPROACH: SWAP ROUNDING

For the algorithm we use to address this problem and create new treatment assignments, we take inspiration from the swap rounding algorithm described by Chekuri and Vondrák, 2009 for B-uniform matroid polytopes. Swap rounding iteratively selects a pair of fractional entries and transfers as much probability mass as possible between them—using a probabilistic randomized decision—to attempt to drive the entries to 0 or 1 until an integral solution is achieved for all units. We outline this algorithm in psuedo-code in Algorithm 1 (full algorithm described in Appendix).

Algorithm 1: Concise Swap Rounding Pseudocode with Explicit Probabilities

Input: A fractional vector p with $\sum_{i=1}^{n} p_i = B$
Output: A binary vector A with $\sum_{i=1}^{n} A_i = B$

```
Set p^(0) ← p, t ← 0
while p^(t) is not binary do
    Select two indices i, j with 0 < p_i^(t), p_j^(t) < 1
    if p_i^(t) + p_j^(t) ≤ 1 then
        Case 1: With probability p_i^(t) / (p_i^(t) + p_j^(t)), set
        p_i^(t+1) ← p_i^(t) + p_j^(t) and p_j^(t+1) ← 0;
        otherwise, set p_i^(t+1) ← 0 and
        p_j^(t+1) ← p_i^(t) + p_j^(t)
    end
    else
        Case 2: With probability (1 − p_i^(t)) / (2 − p_i^(t) − p_j^(t)), round
        p_i^(t) up to 1 (i.e., set p_i^(t+1) ← 1) and let
        p_j^(t+1) ← p_i^(t) + p_j^(t) − 1; otherwise, round
        p_j^(t) up to 1 and let p_i^(t+1) ← p_i^(t) + p_j^(t) − 1
    end
    For all other indices k, set p_k^(t+1) ← p_k^(t)
    Increment t
end
return A ← p^(t)
```

The key features of the swap rounding algorithm are as follows:

- **Exact Feasibility:** Starting with the fractional vector p (with $\sum_{i=1}^{n} p_i = B$), the algorithm iteratively "swaps" probability mass between pairs of units. In each step, it carefully transfers probability so that at the end, every p_i becomes either 0 or 1, ensuring that exactly B units are treated.
- **Marginal Preservation:** The swapping is performed in a manner that guarantees the marginal probability for each unit remains equal to its original value p_i. Thus, even after rounding, we have

$$\Pr(A_i = 1) = p_i \quad \text{for all } i.$$

- **Negative Correlation:** By coupling the rounding decisions where one probability is increased while the other is decreased, swap rounding induces negative correlations between the binary A. We will formally show below that this reduces the variance of treatment effect estimators compared to independent randomization.

In our setting, the use of swap rounding is particularly appealing because it simultaneously satisfies the hard resource constraint and preserves the designed assignment probabilities. This dual property is essential for both fair allocation in real-world interventions and for achieving efficient estimation in statistical analyses.

3.2 NOTATIONS AND ASSUMPTIONS

Each unit has potential outcomes $Y_i(1)$ and $Y_i(0)$ under treatment and non-treatment respectively. We consider two possible sets of distributional assumptions for the units. We start with the design-based setting where potential outcomes are viewed as fixed. A subset of our key results hold in this setting, e.g., Proposition 1 guaranteeing unbiased estimates and Proposition 2 guaranteeing that swap rounding reduces variance. Then, we impose additional assumptions and move to the superpopulation setting where $Y_i(1), Y_i(0)$ are drawn iid from some distribution $\mathbb{P}$. We show that imposing this assumption enables additional results like a central limit theorem and the construction of confidence intervals. In the superpopulation setting, we optionally consider unit-level covariates. That is, for each unit i, we optionally observe covariates V_i where $Y_i(0), Y_i(1), V_i$ are jointly sampled iid from some distribution $\mathbb{P}$ and the assignment probability is a function of the covariates $p(V_i) = \mathbb{P}(A = 1|V_i)$.

Given that this is an RCT, we assume ignorability $(Y(1), Y(0) \perp\!\!\!\perp A)$, and the standard assumptions of consistency ($Y = Y(A)$ whenever $A = a$), and positivity $(p_i, (1 - p_i), p(V_i), 1 - p(V_i) > 0)$. We will also that $Y(0)$ and $Y(1)$ are always nonnegative (this can simply be achieved by scalar translation of Y if it does not hold a-priori).

4 METHODOLOGY

4.1 IPW ESTIMATOR

Our goal in this section is to analyze the statistical properties of the binary treatment assignments produced by our swap rounding procedure. In particular, we focus on the inverse probability weighted (IPW) estimator for the average treatment effect (ATE) when these assignments are used. In later sections, we will further extend our analysis to general linear estimators. Here, we will show that:

- The sequence of partial IPW estimators, constructed using the intermediate fractional values from swap rounding, forms a martingale. This property is crucial for establishing that the final estimator is unbiased.
- By inducing negative correlations among treatment assignments, swap rounding yields an estimator with a lower variance compared to one based on independent assignment.

4.1.1 Treatment effect estimators

Our main results compare two ways of drawing the assignment variables and we introduce separate notation for both. First, A_i is sampled independently with probability p_i. Second, A' is produced via swap rounding applied to $\{p_i\}_{i=1}^n$. The corresponding inverse propensity weighted estimators for the average treatment effect (ATE) τ are

$$\tau_{\text{IPW}} = \frac{1}{n}\sum_{i=1}^{n} \frac{A_i\, Y_i(1)}{p_i} - \frac{(1-A_i)\, Y_i(0)}{1-p_i}, \tag{1}$$

$$\tau_{\text{swap}} = \frac{1}{n}\sum_{i=1}^{n} \frac{A'_i\, Y_i(1)}{p_i} - \frac{(1-A'_i)\, Y_i(0)}{1-p_i}. \tag{2}$$

4.1.2 Martingale Property of the Estimator

Let $p_i^{(0)}$ denote the initial target (fractional) treatment probability for unit i, and let $p_i^{(t)}$ denote the (partial) probability after the tth swap rounding step. Furthermore, let T denote the total number of swap-rounding time steps. Note that $T \le m$ since swap rounding converts at least one of the m entries to an integer at each step and may terminate early if it converts two to integers on a single step (which occurs when $p_i = p_j = 0.5$). Define the IPW-based estimator at step t as

$$x_t = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{p_i^{(t)}\, Y_i(1)}{p_i^{(0)}} - \frac{1-p_i^{(t)}}{1-p_i^{(0)}}\, Y_i(0)\right). \tag{3}$$

Because the swap rounding procedure is constructed so that

$$\mathbb{E}\Big[p_i^{(t)} \,\Big|\, p_i^{(t-1)}\Big] = p_i^{(t-1)}, \tag{4}$$

it follows immediately that

$$\mathbb{E}\Big[x_t \,\Big|\, x_{t-1}\Big] = x_{t-1}. \tag{5}$$

Thus, the sequence $\{x_t\}$ forms a martingale. When $t = T$ (i.e., after complete rounding), $p_i^{(T)} = A_i \in \{0,1\}$ and x_T is the final IPW estimator. This martingale property allows us to establishing unbiasedness as it implies that $\mathbb{E}[X_T] = E[X_0]$. Formally, let us define the Sample Average Treatment Effect (ATE) as $\tau_{SATE} = \frac{1}{n}\sum_{i=0}^{n} Y_i(1) - Y_i(0)$. In the superpopulation setting, we can also define the Population Average Treatment Effect as $\tau_{PATE} = \mathbb{E}[Y(1) - Y(0)]$.

Proposition 1 (Unbiasedness of IPW Swap Rounding for sample ATE)**.** *Let $Y_{1:n}(0), Y_{1:n}(1)$ denote the vectors of sample potential outcomes. Conditional on $Y_{1:n}(0), Y_{1:n}(1)$ (and so including in the design-based setting), we have*

$$\mathbb{E}\left[\tau_{\text{swap}} | Y_{1:n}(0), Y_{1:n}(1)\right] = \tau_{SATE}.$$

In the superpopulation setting, we additionally have $\mathbb{E}[\tau_{\text{swap}}] = \tau_{PATE}$, where the expectation is taken over A and the $Y_{1:n}(0), Y_{1:n}(1)$.

Intuitively, this result holds because swap rounding guarantees that the marginal probability for any individual to be treated is identical as if the assignments were sampled independently, and so the result follows from linearity of expectation. Next, we leverage negative correlation between assignments to show that τ_{swap} can only have lower variance than τ_{IPW}.

4.1.3 Variance Decomposition

Define the individual summand for each unit appear in the IPW estimator as

$$X_i = \frac{A'_i\, Y_i(1)}{p_i^{(0)}} - \frac{(1-A'_i)\, Y_i(0)}{1-p_i^{(0)}}. \tag{6}$$

The variance of the final estimator x_T can be written as

$$\mathbb{V}\big(x_T\big) = \frac{1}{n^2}\left[\sum_{i=1}^{n}\mathbb{V}\big(X_i\big) + 2\sum_{i,j\in S}\text{cov}\big(X_i, X_j\big)\right], \tag{7}$$

where S is the set of T pairs of indices where the swap rounding algorithm made swaps. The key advantage of swap rounding is that it introduces negative correlations between the A'_i's. In particular, if we denote

$$\rho_{ij} = \text{cov}\big(A'_i, A'_j\big), \tag{8}$$

then one can show that

$$\rho_{ij} = \begin{cases} -\,p_i^{(0)}\, p_j^{(0)}, & \text{if } p_i^{(0)} + p_j^{(0)} \le 1, \\ -\big(1-p_i^{(0)}\big)\big(1-p_j^{(0)}\big), & \text{if } p_i^{(0)} + p_j^{(0)} > 1. \end{cases} \tag{9}$$

This negative covariance (which does not occur with independent sampling) decreases the overall variance of x_T relative to the estimator based on independent assignments.

4.1.4 Standard Estimator Variance

From the variance of the IPW estimator (derivation in appendix), we have $\mathbb{V}(X_i|Y_i(0), Y_i(1))$ as

$$\frac{Y_i(1)^2}{p_i^0} + \frac{Y_i(0)^2}{1-p_i^0} - [Y_i(1) - Y_i(0)]^2 \tag{10}$$

4.1.5 Covariance Analysis

Utilizing the fact that the swap rounding procedure preserves the marginals but induces a negative covariance between the two treatment assignments involved in the swap leads (with omitted proof in the Appendix) to the final covariance expression: $\text{cov}(X_i, X_j|Y_{i,j}(0), Y_{i,j}(1)) =$

$$\rho_{ij}\left(\frac{Y_i(1)}{p_i^{(0)}} + \frac{Y_i(0)}{1-p_i^{(0)}}\right)\left(\frac{Y_j(1)}{p_j^{(0)}} + \frac{Y_j(0)}{1-p_j^{(0)}}\right) \tag{11}$$

Because of our non-negativity assumption on outcome values and the fact that ρ_{ij} is negative, we see that $\text{cov}(X_i, X_j)$ is always negative as ρ_{ij} remains negative. Thus, by "coupling" the rounding of different units, our procedure further reduces variance relative to independent assignment. As 4.4 details the standard variance of the IPW variance, the fact that $\text{cov}(X_i, X_j)$ is always negative means that we have Proposition 2.

Proposition 2 (Variance Reduction)**.** *Under the swap rounding procedure, the negative dependence between the assignment indicators ensures that for any value of the potential outcomes $Y_{i:n}(0), Y_{i:n}(1)$*

$$\mathbb{V}\left(\tau_{\text{swap}}|Y_{i:n}(0), Y_{i:n}(1)\right) \leq \mathbb{V}\left(\tau_{\text{IPW}}|Y_{i:n}(0), Y_{i:n}(1)\right).$$

If there exists at least one swapped pair i, j where $Y_i(0), Y_i(1), Y_j(0), Y_j(1)$ are not all simultaneously zero, the inequality is strict. In the superpopulation setting, we also have unconditionally that $\mathbb{V}\left(\tau_{\text{swap}}\right) \leq \mathbb{V}\left(\tau_{\text{IPW}}\right)$.

This indicates that precision of inference will never decrease if we draw the assignments using swap rounding.

4.1.6 Inference

We now turn to providing inferential guarantees, including asymptotically valid confidence intervals, for the swap rounding estimator. Our first task is to construct a consistent estimator for $\mathbb{V}\left(\tau_{\text{swap}}\right)$. At this stage, all of our results will require the superpopulation assumption that units are sampled iid. The challenge is that terms of the form $[Y_i(1) - Y_i(0)]^2$ appear in the variance of the estimator conditional on the potential outcomes, but both potential outcomes are never observed simultaneously for the same unit. In the superpopulation setting, the unconditional variance of the estimator instead contains terms of the form $\mathbb{E}[Y(1) - Y(0)]^2$ that can be proxied with τ_{swap} itself. For all other terms, appearing in $\mathbb{V}\left(\tau_{\text{swap}}\right)$, we use inverse-propensity weighted estimates. Let $Y_i = Y_i(A_i')$ denote the observed outcome for unit i. This yields the final variance estimator:

$$\hat{\sigma} = \frac{1}{n^2}\Bigg[\sum_{i=1}^{n}\Big(\frac{A_i'(Y_i)^2}{(p_i^0)^2} + \frac{(1-A_i')(Y_i)^2}{(1-p_i^0)^2}\Big) - n\tau_{\text{swap}}^2 \tag{12}$$

$$+ 2\sum_{i,j\in S}\rho_{ij}\left(\frac{A_i'Y_i}{p_i^{(0)}} + \frac{(1-A_i')Y_i}{1-p_i^{(0)}}\right) \tag{13}$$

$$\left(\frac{A_j'Y_j}{p_j^{(0)}} + \frac{(1-A_j')Y_j}{1-p_j^{(0)}}\right)\Bigg] \tag{14}$$

As the assignment probabilities p_i are known exactly, we establish that:

Proposition 3 (Unbiasedness and Consistency of IPW Swap Rounding Variance Estimator)**.** *$\hat{\sigma}$ is a consistent and unbiased Estimator of $\mathbb{V}[\tau_{\text{swap}}]$ under the superpopulation assumption.*

Finally, we use a martingale central limit theorem to show that τ_{swap} is asymptotically normal. Combined with Proposition 3, this enables the construction of valid confidence intervals.

Theorem 1 (Asymptotic Normality and Valid Confidence Intervals)**.** *Given the superpopulation assumption and under the further assumptions that:,*

- ***Bounded Moments:*** *$Y_i(1)$ and $Y_i(0)$ have uniformly bounded second moments by a constant $M < \infty$.*
- ***Lower Bound on Quadratic Variation:*** *There exists a constant $c > 0$ such that for all sufficiently large n,*

$$\langle x\rangle_T \;=\; \frac{1}{n}\sum_{t=1}^{T}\mathbb{E}\big[\Delta_t^2 \mid x_{t-1}\big] \;\geq\; c\,.$$

Them, the IPW estimator obtained via swap rounding, τ_{swap}, is asymptotically normal. That is,

$$\sqrt{n}\left(\tau_{\text{swap}} - \tau\right) \xrightarrow{d} N(0, \sigma^2),$$

where σ^2 is the asymptotic variance given by the expression in Equation 7. Consequently, we have for $\widehat{\sigma}^2$, the $1-\alpha$

confidence interval

$$\tau_{\text{swap}} \pm z_{1-\alpha/2} \frac{\widehat{\sigma}}{\sqrt{n}}$$

has asymptotically nominal coverage.

4.2 GENERAL PROPERTIES OF LINEAR ESTIMATORS UNDER SWAP ROUNDING

In fact, these properties extend beyond the IPW estimator as we show that a key property of swap rounding is that any estimator that is linear in the treatment assignments is automatically unbiased and has lower variance relative to the case of independent assignments. To state this formally, we have Theorem 2 (omitted proof in Appendix).

Theorem 2 (Unbiasedness and Variance Reduction for Linear Estimators)**.** *Let*

$$\psi = \sum_{i=1}^{n} a_i A_i + b \tag{15}$$

be an estimator that is linear in the treatment assignments A_i for $a_i \geq 0$. Define ψ_{swap} and ψ_{IPW} as the versions of this estimator where the A are sampled from swap rounding and independently, respectively. Treating a_i and b as fixed and analyzing the estimators conditional on them,

$$\mathbb{E}[\psi_{swap}] = \mathbb{E}[\psi_{IPW}] = \sum_{i=1}^{n} a_i\, p_i + b, \tag{16}$$

$$\begin{aligned}\mathbb{V}(\psi_{swap}) &= \sum_{i=1}^{n} a_i^2\, \mathbb{V}(A_i') + 2 \sum_{i,j \in S} a_i\, a_j\, cov(A_i', A_j') \\ &\leq \sum_{i=1}^{n} a_i^2\, \mathbb{V}(A_i) = \mathbb{V}(\psi_{IPW}),\end{aligned} \tag{17}$$

i.e., the variance is lower than (or equal to) what would be achieved under independent assignment. If there is any pair where $a_i, a_j \neq 0$, the inequality is strict.

5 COVARIATE-ORDERED SWAP ROUNDING

In addition to preserving marginal treatment probabilities and inducing negative correlations through swap rounding, it is beneficial to pair units with similar covariates to further reduce variance. The intuition is that when units are similar in their covariate profiles, they are likely to have similar baseline outcomes. Thus, any difference induced by the randomized rounding is less variable. This idea parallels matching methods in causal inference, where pairing or matching treated and control units based on covariate similarity improves the precision of effect estimates [Stuart, 2010]. We will introduce an extension of our swap-rounding methodology that leverages this intuition.

We first order the n units by their covariate vectors V_i using a distance metric

$$d(V_i, V_j) = \|V_i - V_j\|,$$

and solve

$$\min_{\pi \in \mathcal{P}(n)} \sum_{k=2}^{n} d\Big(V_{\pi(k)}, V_{\pi(k-1)}\Big), \tag{18}$$

where $\mathcal{P}(n)$ denotes the set of all permutations of $\{1, \ldots, n\}$. (This is analogous to a TSP, which can be approximately solved via heuristics Johnson and McGeoch [1997], Kirkpatrick et al. [1983].) Swap rounding is then applied only to adjacent pairs in the ordering, ensuring that negative correlations are induced primarily between similar units.

Next, we layout the assumptions needed to ensure covariate ordering swapping reduces variance. While these assumptions are restrictive, this should be seen a stylized model to get a sense of ideal conditions that would exist for covariate ordering to be beneficial, and not something we expect to always hold. Essentially, we want to show that this technique works better for conditions when outcomes are closer for similar covariates than for dissimilar covariates.

Assume the potential outcomes follow

$$Y_i(0) = f(V_i) + \epsilon_{i0}, \tag{19}$$

$$Y_i(1) = f(V_i) + \tau(V_i) + \epsilon_{i1}, \tag{20}$$

with f and τ Lipschitz (with constants L_f and L_τ, respectively), ϵ as random mean-zero error terms that are independent of the V, and that the target treatment probabilities $p_i^{(0)}$ are Lipschitz in V_i (with constant L_p).

Let A_i' be the final assignment from swap rounding and

$$\text{Cov}(A_i', A_j') = \rho_{ij} < 0 \quad (i \neq j).$$

The unit-level IPW estimator is

$$X_i = \frac{A_i'\, Y_i(1)}{p_i^{(0)}} - \frac{(1 - A_i')\, Y_i(0)}{1 - p_i^{(0)}}.$$

Define the effective weight by

$$\begin{aligned}M_i &\triangleq \frac{\mathbb{E}[Y_i(1)]}{p_i^{(0)}} + \frac{\mathbb{E}[Y_i(0)]}{1 - p_i^{(0)}} \\ &= \frac{f(V_i) + \tau(V_i)}{p_i^{(0)}} + \frac{f(V_i)}{1 - p_i^{(0)}}.\end{aligned} \tag{21}$$

Then, it can be shown that

$$\text{Cov}(X_i, X_j) = \rho_{ij}\, M_i\, M_j.$$

To exploit covariate similarity, we assume that $M(\cdot)$ is *bi-Lipschitz*; i.e., there exist constants $\ell_M, L_M > 0$ such that for all V_i, V_j

$$\ell_M \|V_i - V_j\| \leq |M_i - M_j| \leq L_M \|V_i - V_j\|. \quad (22)$$

Let S_{opt} denote the set of pairs of neighboring units that are swapped under the covariate-optimized pairing and S_{rand} be a fixed alternative ordering. We will compare the variance of treatment effect estimates for the swap rounding estimator under S_{opt} to S_{rand}. This will require formalizing the idea that the optimization algorithm succeeds in finding an ordering where neighboring units are all significantly closer than the average under S_{rand}. First, for S_{opt}, suppose that

$$\|V_i - V_j\| \leq \delta,$$

for all i, j for some $\delta > 0$. Hence,

$$|M_i - M_j| \leq L_M \, \delta \triangleq \Delta_{\text{opt}}. \quad (23)$$

In contrast, let the expected squared covariate distance under S_{rand} be

$$\mathbb{E}_{(i,j)\sim S_{\text{rand}}}\left[\|V_i - V_j\|^2\right] = c^2, \quad (24)$$

for some $c > 0$. Now, assume that $c > \delta$, i.e., that the optimized ordering places all neighboring units closer than the average under . Then by the lower bound in (22) and Jensen's inequality,

$$\Delta_{\text{rand}}^2 \triangleq \mathbb{E}_{(i,j)\sim S_{\text{rand}}}\left[(M_i - M_j)^2\right] \geq \ell_M^2 \, c^2,$$
$$\Delta_{\text{rand}} \geq \ell_M \, c. \quad (25)$$

Thus, if

$$L_M \, \delta < \ell_M \, c, \quad (26)$$

then $\Delta_{\text{opt}} < \Delta_{\text{rand}}$. Since

$$M_i M_j = \left(\frac{M_i + M_j}{2}\right)^2 - \left(\frac{M_i - M_j}{2}\right)^2,$$

a smaller $|M_i - M_j|$ (for a fixed average) leads to a larger $M_i M_j$. Consequently, the negative covariance $\rho_{ij} \, M_i M_j$ is stronger under covariate-ordered pairing, reducing the overall variance of the IPW estimator.

Proposition 4 (Variance Reduction via Covariate-Ordered Pairing)**.** *Assume that units are paired via covariate-ordered swap rounding so that for each pair (i, j) in the optimal pairing S_{opt} we have*

$$\|V_i - V_j\| \leq \delta,$$

and that the effective weight function $M(\cdot)$ satisfies (22)*. Further, suppose an arbitrary pairing S_{rand} satisfies*

$$\mathbb{E}_{(i,j)\sim S_{\text{rand}}}\left[\|V_i - V_j\|^2\right] \geq c^2, \quad (27)$$

with $c > \delta$ and that condition (26) *holds. Then,*

$$\mathbb{V}\big(\tau_{\text{IPW}}\big)_{S_{\text{opt}}} < \mathbb{V}\big(\tau_{\text{IPW}}\big)_{S_{\text{rand}}}.$$

6 EXPERIMENTS

6.1 GENERAL SETUP

For our experiments, we generate consider 3 different settings. Firstly, a fully synthetic setting where we generate potential outcomes, assignment probabilities and covariates through simulation. We then consider a semi-synthetic setting involving 2 real datasets.

6.2 BASELINES

Below, we briefly describe the baseline methods we compare our swap-rounding IPW estimator and Covariate-Matched Swap-Rounding IPW Estimator to.

1. **Standard IPW Estimator** We use the standard construct of the Horvitz–Thompson IPW estimator
2. **Budget Limited Bernoulli Assignment:** A conditional assignment procedure that repeatedly draws independent Bernoulli assignments until the draw satisfies $\sum_i A_i = B$, akin to conditional Poisson sampling Hájek [1964], Grafström [2005].
3. **Re-randomization with Effective Propensity Adjustment:** Multiple candidate assignments are drawn from the Bernoulli model, and one is selected based on a Mahalanobis distance criterion that enforces covariate balance (see Morgan and Rubin [2012]). An effective propensity is then computed using the methodology described in Branson and Miao [2019] for max sampling for optimal selection and used to adjust the IPW estimator to un-bias the estimates.
4. **Random Assignment Based on Total Probabilities:** Exactly B units are selected uniformly at random (without replacement) for treatment. This corresponds to simple random sampling Cochran [1977]. As such, this method does not preserve target probability values p^0.
5. **Self-Normalized IPW:** With independent Bernoulli assignments, the self-normalized IPW estimator is computed as

$$\hat{\tau} = \frac{\sum_i A_i Y_i / p_i}{\sum_i A_i / p_i} - \frac{\sum_i (1 - A_i) Y_i / (1 - p_i)}{\sum_i (1 - A_i) / (1 - p_i)},$$

This method was designed to reduce the impact of extreme weights and to reduced variance (which it tends to do extremely well) as compared to standard IPW estimators Liu et al. [2016]. However, in finite samples this estimator is not unbiased. [Datta and Polson, 2011]. While this method and Random Assignment are biased, we include them to show the variance we hope to empirically achieve.

6.3 SYNTHETIC EXPERIMENTS

6.3.1 Synthetic Experiment Setup

For each sample size n (we use various sample sizes), we generate 100 unique scenarios with covariates $V \in \mathbb{R}^{n\times 3}$ (with entries drawn from a standard normal), and outcomes given by

$$Y_i(0) = \beta_0 + V_i^\top \beta, \quad Y_i(1) = Y_i(0) + \tau_{\text{true}} + \epsilon_i,$$

with $\tau_{\text{true}} = 2$ and $\epsilon_i \sim \mathcal{N}(0,1)$. The distribution of the initial probabilities p^0 is computed via the methodology described in section 6.1.

In this setting, we consider 3 methodologies for assigning target treatment probabilities $p^{(0)}$.

1. **Uniform Distribution:** $p^{(0)} \sim$ Uniform (0.01,0.99)
2. **Gaussian Distribution:** $p^{(0)} \sim \mathcal{N}(.5, .25)$ with probabilities clipped between (0.01,0.99)
3. **Covariate Based Distribution:** p_0 is computed via a logistic model of V (clipping probabilities between .01 and .99)

Each of these scenarios is then replicated 100 times. All code is provided is supplementary materials.

6.3.2 Experiments

We conduct 3 separate experiments, each one different by the methodology of how the target probabilities $p^{(0)}$ are distributed. In all 3 experiments (Uniform, Gaussian and Covariate-Based Distribution seen in Figures 1, 2, 3), we notice a general trend. Covariate Based Swap Rounding performs the best (minimum variance) by a considerable margin, followed by Random Assignment, then standard Swap Rounding, Self-Normalized IPW and then other methods. This result is especially significant because our methodology outperforms estimators like Self-Normalized IPW and Random Assignment that were designed to minimize variance at the cost of bias. We also notice several other interesting trends such as the standard IPW estimator performing worst in all experiments. At large sample sizes, all estimators converge to negligible variance.

6.4 SEMI-SYNTHETIC DATA

For the type of data that is needed to test these methods, we need access to the true propensity scores underlying trial selection as well as access to both potential outcomes for each individual. This obviously is not possible so we have to sacrifice data quality in some places and make it up with simulation. That being said, we proceed with two data sources that provide different levels of information.

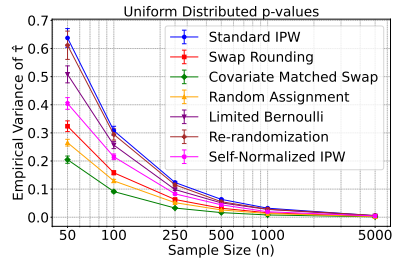

Figure 1: Synthetic Experiment with $p^{(0)} \sim$ Uniform (0.01,0.99). Mean empirical variance with 95 % CI.

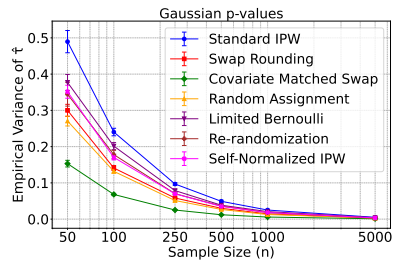

Figure 2: Synthetic Experiment with $p^{(0)} \sim \mathcal{N}(.5, .25)$

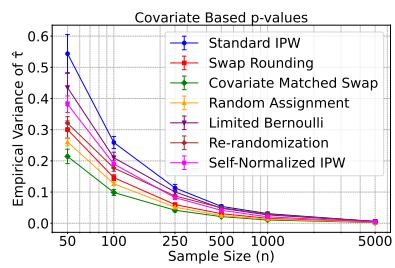

Figure 3: Synthetic Experiment where p_0 is computed via a logistic model of the covariates V

6.4.1 IDHP Dataset

Our first data source is the IDHP Dataset. [Gross, 2024] The dataset (747 individuals) is based off an RCT that was conducted between 1985-1988 where they tested the impact of early interventions in reducing developmental problems associated with low weight and premature infants where the outcome was a continuous measure of cognitive development. We used simulated counterfactual outcomes for each individual as well propensity scores estimated via logistic regression. Our results are shown in Figure 4 (the Random Assignment line is hidden underneath the Self-Normalized IPW line). In this setting again our methodology remains competitive with the biased methodologies while performing superior to those methods who are unbiased.

6.4.2 Housing Dataset

In this setting, we use data provided by the Allegheny County Department of Human Services. The dataset (3163

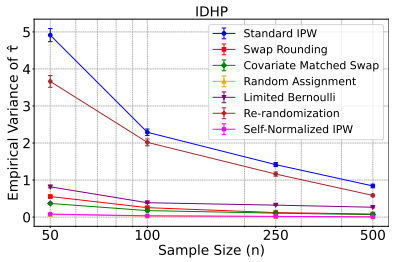

Figure 4: Semi-Synthetic Experiment IDHP dataset

individuals) was introduced by Wilder and Welle [2024] and contains individuals eligible for public housing assistance and simulates the effect of an intervention on the binary outcome whether they will have at least 4 emergency room visits in the upcoming year. The hypothetical intervention's effect is characterized by $\tau(X) = .1 \cdot P(Y(0) = 1 \mid X)$ where 0.1 quantifies the proportional decrease in the occurrence of an adverse outcomes resulting from the treatment, and the probability $P(Y(0) = 1 \mid X)$ is estimated using a random forest model. Our results are shown in Figure 5. In this experiment, regular swap rounding performs the best.

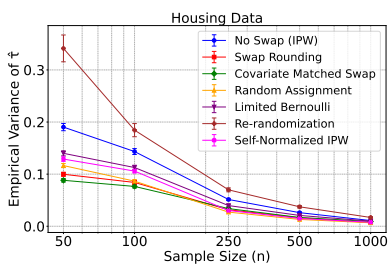

Figure 5: Semi-Synthetic Experiment Housing Dataset

7 DISCUSSION

We theoretically show that incorporating swap rounding, due the negative correlations it induces, leads to an unbiased and variance-reduced IPW estimator that can enforce strict budget constraints. Moreover, by incorporating covariate-ordered pairing, our approach further refines variance reduction, particularly in settings where units share similar characteristics. These findings highlight the potential of our method for designing efficient and fair experiments in resource-constrained environments.

8 ACKNOWLEDGMENTS

This project was supported by the AI2050 program at Schmidt Sciences (Grant G-22-64474) and the AI Research Institutes Program funded by the National Science Foundation under AI Institute for Societal Decision Making (AI-SDM), Award No. 2229881. We would also like to acknowledge Naveen Raman, Nima Jadali and Albert Ting for their contributions in editing and advising on the paper.

References

A. A. Ageev and M. Sviridenko. Approximation algorithms for maximization of submodular functions. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing (STOC)*, pages 793–802, 2004.

Nibia Aires. Comparisons between conditional poisson sampling and pareto πps sampling designs. *Journal of Statistical Planning and Inference*, 88(1):133–147, 2000. ISSN 0378-3758. doi: https://doi.org/10.1016/S0378-3758(99)00205-0. URL https://www.sciencedirect.com/science/article/pii/S0378375899002050.

S. Athey and G. W. Imbens. The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, 31(2):3–32, 2017.

R Branson and Q Miao. Maximally balanced sampling in randomized experiments. *Journal of Statistical Planning and Inference*, 205:1–10, 2019.

Sylvain Chassang, Gerard Padró I Miquel, and Erik Snowberg. Selective trials: A principal-agent approach to randomized controlled experiments. *American Economic Review*, 102(4):1279–1309, June 2012. doi: 10.1257/aer.102.4.1279. URL https://www.aeaweb.org/articles?id=10.1257/aer.102.4.1279.

Chandra Chekuri and Jan Vondrák. Randomized pipage rounding for matroid polytopes and applications. *CoRR*, abs/0909.4348, 2009. URL http://arxiv.org/abs/0909.4348.

William G Cochran. *Sampling Techniques*. John Wiley & Sons, 1977.

Jyotishka Datta and Nicholas G. Polson. Inverse probability weighting: from survey sampling to evidence estimation. Working paper, Department of Statistics, Virginia Tech and
Booth School of Business, University of Chicago, 2011. Unpublished manuscript.

Anton Grafström. Comparisons of methods for generating conditional poisson samples and sampford samples. 2005. URL https://api.semanticscholar.org/CorpusID:56873261.

Ruth T. Gross. Infant health and development program (ihdp): Enhancing the outcomes of low birth weight, premature infants in the united states, 1985–1988. Inter-university Consortium for Political and Social Research, 2024. URL https://doi.org/10.3886/

ICPSR09795.v2. Dataset, Version V2, released February 14, 2024.

J Hájek. Asymptotic theory of rejective sampling with varying probabilities. *The Annals of Mathematical Statistics*, pages 1006–1018, 1964.

G. W. Imbens and D. B. Rubin. *Causal Inference in Statistics, Social, and Biomedical Sciences*. Cambridge University Press, 2015.

C. Isaki and W. A. Fuller. Estimation for finite population sampling. *Journal of the American Statistical Association*, 79(387):137–145, 1984.

David S. Johnson and Laurence A. McGeoch. The traveling salesman problem: A case study in local optimization. In *Local Search in Combinatorial Optimization*, pages 215–310. Princeton University Press, 1997.

Daido Kido. Incorporating preferences into treatment assignment problems, 2023. URL https://arxiv.org/abs/2311.08963.

S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

X. Li, P. Ding, and D. B. Rubin. Rerandomization in experiments: A review. *Annual Review of Statistics and Its Application*, 5:1–19, 2018.

Lan Liu, Michael G Hudgens, and Sara Becker-Dreps. On inverse probability-weighted estimators in the presence of interference. *Biometrika*, 103(4):829–842, 2016. doi: 10.1093/biomet/asw047.

D. L. Morgan and D. B. Rubin. Rerandomization to improve covariate balance in experiments. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74 (4):515–532, 2012.

Yusuke Narita. Incorporating ethics and welfare into randomized experiments. *Proceedings of the National Academy of Sciences*, 118(1):e2008740118, 2021. doi: 10.1073/pnas.2008740118. URL https://www.pnas.org/doi/abs/10.1073/pnas.2008740118.

J. Neyman. On the problem of the most efficient allocation of units in stratified sampling. *Journal of the Royal Statistical Society*, 97:544–557, 1934.

A. Srinivasan. Distributions on level-sets with applications to approximation algorithms. In *Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 588–597, 2001.

Elizabeth A. Stuart. Matching methods for causal inference: A review and a look forward. *Statistical Science*, 25(1): 1–21, 2010.

Bryan Wilder and Pim Welle. Learning treatment effects while treating those in need, 2024. URL https://arxiv.org/abs/2407.07596.

9 SWAP ROUNDING ALGORITHM

Algorithm 2: Simplified Swap Rounding for B-Uniform Matroid

Input: Fractional vector $p \in [0,1]^n$ with $\sum_{i=1}^n p_i = B$
Output: Binary vector $A \in \{0,1\}^n$ with $\sum_{i=1}^n A_i = B$

$t \leftarrow 0,\ p^{(0)} \leftarrow p$
while $p^{(t)}$ *is not integral,* **do**
  Select indices i and j such that $0 < p_i^{(t)} < 1$ and $0 < p_j^{(t)} < 1$
  if $p_i^{(t)} + p_j^{(t)} \leq 1$ **then**
    Let $\delta = \min\{p_i^{(t)}, p_j^{(t)}\}$
    With probability $\frac{p_i^{(t)}}{p_i^{(t)}+p_j^{(t)}}$, set

$$p_i^{(t+1)} = p_i^{(t)} - \delta, \quad p_j^{(t+1)} = p_j^{(t)} + \delta;$$

    otherwise set

$$p_i^{(t+1)} = p_i^{(t)} + \delta, \quad p_j^{(t+1)} = p_j^{(t)} - \delta;$$

  else
    Let $\delta = \min\{1 - p_i^{(t)}, 1 - p_j^{(t)}\}$
    With probability $\frac{1-p_i^{(t)}}{2-p_i^{(t)}-p_j^{(t)}}$, set

$$p_i^{(t+1)} = p_i^{(t)} + \delta, \quad p_j^{(t+1)} = p_j^{(t)} - \delta;$$

    otherwise set

$$p_i^{(t+1)} = p_i^{(t)} - \delta, \quad p_j^{(t+1)} = p_j^{(t)} + \delta;$$

  end
  For all $k \notin \{i,j\}$, let $p_k^{(t+1)} = p_k^{(t)}$
  $t \leftarrow t+1$
end
return $A = p^{(t)}$

10 IPW PROOFS

10.1 IPSW- SWAP ROUNDING

We want to figure out which subjects to assign treatments to so that we can use an IPSW estimator to estimate $\hat{\tau}$ - the average treatment effect. We start with a vector of probabilities p. This vector then undergoes iterative swap rounding results in A, a vector of binary treatments. The negative correlation aspect of swap rounding allows us to maintain concentration bounds when we don't have a very large number of samples. We observe the outcomes Y(A) and use them in our IPSW estimator for $\hat{\tau}$. We define the following variables:
p_t: probabilities after step t of swap rounding
p_0: target probabilites (pre-rounding)
p_T: $A_1, ..., A_n$ (binary treatments)

$$x_t = \frac{1}{n}\sum_i \frac{p_i^t Y_i(1)}{p_i^0} - \frac{(1-p_i^t)Y_i(0)}{1-p_i^0}$$
$$x_0 = \frac{1}{n}\sum_i Y(1) - Y(0)$$
$$x_T = \frac{1}{n}\sum_i \frac{A_i Y(1)}{p_i^0} - \frac{(1-A_i)Y(0)}{1-p_i^0}$$

10.2 SHOWING MARTINGALE PROPERTY

First, we want to show that

$$\mathbb{E}[x_t|x_{t-1}] = x_{t-1}$$

We have

$$\begin{aligned}\mathbb{E}[x_t|x_{t-1}] &= \frac{1}{n}\sum_i \frac{\mathbb{E}[p_i^t|p_i^{t-1}]Y_i(1)}{p_i^0} - \frac{(1-\mathbb{E}[p_i^t|p_i^{t-1}])Y_i(0)}{1-p_i^0} \\ &= \frac{1}{n}\sum_i \frac{p_i^{t-1}Y_i(1)}{p_i^0} - \frac{(1-p_i^{t-1})Y_i(0)}{1-p_i^0} \\ &= x_{t-1}\end{aligned}$$

10.3 VARIANCE-COVARIANCE DECOMPOSITION

$$\mathbb{V}[X_T] = \frac{1}{n^2}\Big(\sum_{i=1}^{N}\mathbb{V}\Big(\frac{A_iY_i(1)}{p_i^0} - \frac{(1-A_i)Y_i(0)}{1-p_i^0}\Big) + 2\sum_{i,j\in S}\text{Cov}\Big(\frac{A_iY_i(1)}{p_i^0} - \frac{(1-A_i)Y_i(0)}{1-p_i^0}, \frac{A_jY_j(1)}{p_j^0} - \frac{(1-A_j)Y_j(0)}{1-p_j^0}\Big)\Big)$$

where S is the set of swaps in order. For the variance component, we have

SINGLE-UNIT VARIANCE

Define

$$X_i = \frac{A_i' Y_i(1)}{p_i^0} - \frac{(1-A_i')Y_i(0)}{1-p_i^0}.$$

Then

$$\begin{aligned}X_i^2 &= \left(\frac{A_i' Y_i(1)}{p_i^0} - \frac{(1-A_i')Y_i(0)}{1-p_i^0}\right)^2 \\ &= \frac{A_i' Y_i(1)^2}{(p_i^0)^2} + \frac{(1-A_i')Y_i(0)^2}{(1-p_i^0)^2} - 2\frac{A_i'(1-A_i')Y_i(1)Y_i(0)}{p_i^0(1-p_i^0)}.\end{aligned}$$

Taking expectations (using $\mathbb{E}[A_i'] = p_i^0$ and $\mathbb{E}[A_i'(1-A_i')] = 0$) yields

$$\mathbb{E}[X_i^2] \;=\; \frac{Y_i(1)^2}{p_i^0} \;+\; \frac{Y_i(0)^2}{1-p_i^0}.$$

Since $\mathbb{E}[X_i] = Y_i(1) - Y_i(0)$, it follows that

$$\mathrm{Var}(X_i) \;=\; \frac{Y_i(1)^2}{p_i^0} \;+\; \frac{Y_i(0)^2}{1-p_i^0} \;-\; \big[Y_i(1) - Y_i(0)\big]^2.$$

SUPER POPULATION

If $\big(Y_i(0),\, Y_i(1),\, p_i^0\big)$ is drawn i.i.d. from a larger (hypothetically infinite) population, then taking an outer expectation of the above result gives

$$\mathbb{V}(X_i) \;=\; \frac{\mathbb{E}\big[Y_i(1)^2\big]}{p_i^0} \;+\; \frac{\mathbb{E}\big[Y_i(0)^2\big]}{1-p_i^0} \;-\; \mathbb{E}\big[Y_i(1) - Y_i(0)\big]^2.$$

10.4 COVARIANCE OF X_i AND X_j

Now, we aim to calculate the covariance - for a pair of units X_i, X_j

$$\mathrm{Cov}(X_i, X_j) = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j],$$

where

$$X_i = \frac{A_i Y_i(1)}{p_i^0} - \frac{(1-A_i)Y_i(0)}{1-p_i^0}, \quad X_j = \frac{A_j Y_j(1)}{p_j^0} - \frac{(1-A_j)Y_j(0)}{1-p_j^0}.$$

10.5 EXPANDING $X_i X_j$

The product $X_i X_j$ expands as:

$$\begin{aligned} X_i X_j &= \left(\frac{A_i Y_i(1)}{p_i^0} - \frac{(1-A_i)Y_i(0)}{1-p_i^0}\right)\left(\frac{A_j Y_j(1)}{p_j^0} - \frac{(1-A_j)Y_j(0)}{1-p_j^0}\right) \\ &= \frac{A_i A_j Y_i(1) Y_j(1)}{p_i^0 p_j^0} - \frac{A_i(1-A_j)Y_i(1)Y_j(0)}{p_i^0(1-p_j^0)} - \frac{(1-A_i)A_j Y_i(0) Y_j(1)}{(1-p_i^0)p_j^0} \\ &\quad + \frac{(1-A_i)(1-A_j)Y_i(0)Y_j(0)}{(1-p_i^0)(1-p_j^0)}. \end{aligned}$$

10.6 TAKING EXPECTATIONS

First Term: $\mathbb{E}\left[\frac{A_i A_j Y_i(1) Y_j(1)}{p_i^0 p_j^0}\right]$

Define $\rho_{ij} = \mathrm{Cov}(A_i, A_j)$. We then have

$$\mathbb{E}[A_i A_j] = p_i^0 p_j^0 + \rho_{ij}.$$

Thus:

$$\mathbb{E}\left[\frac{A_i A_j Y_i(1) Y_j(1)}{p_i^0 p_j^0}\right] = \frac{(p_i^0 p_j^0 + \rho_{ij}) \cdot \mathbb{E}[Y_i(1)] \cdot \mathbb{E}[Y_j(1)]}{p_i^0 p_j^0}.$$

Second Term: $\mathbb{E}\left[\frac{A_i(1-A_j)Y_i(1)Y_j(0)}{p_i^0(1-p_j^0)}\right]$

$$\mathbb{E}[A_i(1-A_j)] = p_i^0(1-p_j^0) - \rho_{ij}.$$

Thus:

$$\mathbb{E}\left[\frac{A_i(1-A_j)Y_i(1)Y_j(0)}{p_i^0(1-p_j^0)}\right] = \frac{(p_i^0(1-p_j^0) - \rho_{ij}) \cdot \mathbb{E}[Y_i(1)] \cdot \mathbb{E}[Y_j(0)]}{p_i^0(1-p_j^0)}.$$

Third Term: $\mathbb{E}\left[\frac{(1-A_i)A_jY_i(0)Y_j(1)}{(1-p_i^0)p_j^0}\right]$

$$\mathbb{E}[(1-A_i)A_j] = (1-p_i^0)p_j^0 - \rho_{ij}.$$

Thus:

$$\mathbb{E}\left[\frac{(1-A_i)A_jY_i(0)Y_j(1)}{(1-p_i^0)p_j^0}\right] = \frac{((1-p_i^0)p_j^0 - \rho_{ij}) \cdot \mathbb{E}[Y_i(0)] \cdot \mathbb{E}[Y_j(1)]}{(1-p_i^0)p_j^0}.$$

Fourth Term: $\mathbb{E}\left[\frac{(1-A_i)(1-A_j)Y_i(0)Y_j(0)}{(1-p_i^0)(1-p_j^0)}\right]$

$$\mathbb{E}[(1-A_i)(1-A_j)] = (1-p_i^0)(1-p_j^0) + \rho_{ij}.$$

Thus:

$$\mathbb{E}\left[\frac{(1-A_i)(1-A_j)Y_i(0)Y_j(0)}{(1-p_i^0)(1-p_j^0)}\right] = \frac{((1-p_i^0)(1-p_j^0) + \rho_{ij}) \cdot \mathbb{E}[Y_i(0)] \cdot \mathbb{E}[Y_j(0)]}{(1-p_i^0)(1-p_j^0)}.$$

10.7 SUBTRACTING $\mathbb{E}[X_i]\mathbb{E}[X_j]$

The product $\mathbb{E}[X_i]\mathbb{E}[X_j]$ is:

$$\mathbb{E}[X_i]\mathbb{E}[X_j] = \left(\frac{p_i^0 \cdot \mathbb{E}[Y_i(1)]}{p_i^0} - \frac{(1-p_i^0) \cdot \mathbb{E}[Y_i(0)]}{1-p_i^0}\right) \cdot \left(\frac{p_j^0 \cdot \mathbb{E}[Y_j(1)]}{p_j^0} - \frac{(1-p_j^0) \cdot \mathbb{E}[Y_j(0)]}{1-p_j^0}\right).$$

After subtracting from $\mathbb{E}[X_iX_j]$, we get

10.8 FINAL COVARIANCE EXPRESSION

The covariance $\text{Cov}(X_i, X_j)$ is:

$$\begin{aligned}\text{Cov}(X_i, X_j) = \rho_{ij} \cdot \Bigg(&\frac{\mathbb{E}[Y_i(1)] \cdot \mathbb{E}[Y_j(1)]}{p_i^0p_j^0} + \frac{\mathbb{E}[Y_i(1)] \cdot \mathbb{E}[Y_j(0)]}{p_i^0(1-p_j^0)} \\ &+ \frac{\mathbb{E}[Y_i(0)] \cdot \mathbb{E}[Y_j(1)]}{(1-p_i^0)p_j^0} + \frac{\mathbb{E}[Y_i(0)] \cdot \mathbb{E}[Y_j(0)]}{(1-p_i^0)(1-p_j^0)}\Bigg).\end{aligned}$$

where

$$\rho_{ij} = \text{Cov}(A_i, A_j) = \mathbb{E}[A_iA_j] - p_i^0p_j^0$$

We then have

$$\mathbb{E}[A_iA_j] = \begin{cases} 0 & \text{if } p_i^0 + p_j^0 \leq 1, \\ p_i^0 + p_j^0 - 1 & \text{if } p_i^0 + p_j^0 > 1. \end{cases}$$

so

$$\rho_{ij} = \begin{cases} -p_i^0 p_j^0 & \text{if } p_i^0 + p_j^0 \leq 1, \\ -(1-p_i^0)(1-p_j^0) & \text{if } p_i^0 + p_j^0 > 1. \end{cases}$$

10.9 VARIANCE OF X_T EXPRESSION

$$\begin{aligned}\mathbb{V}[X_T] = \frac{1}{n^2}\Bigg[\sum_{i=1}^{N}\Big(\frac{\mathbb{E}[Y_i(1)^2]}{p_i^0} + \frac{\mathbb{E}[Y_i(0)^2]}{1-p_i^0}\Big) - n\mathbb{E}[Y(1)-Y(0)]^2 \\ + 2\sum_{i,j\in S}\rho_{ij}\cdot\Big(\frac{\mathbb{E}[Y_i(1)]\cdot\mathbb{E}[Y_j(1)]}{p_i^0 p_j^0} - \frac{\mathbb{E}[Y_i(1)]\cdot\mathbb{E}[Y_j(0)]}{p_i^0(1-p_j^0)} \\ - \frac{\mathbb{E}[Y_i(0)]\cdot\mathbb{E}[Y_j(1)]}{(1-p_i^0)p_j^0} + \frac{\mathbb{E}[Y_i(0)]\cdot\mathbb{E}[Y_j(0)]}{(1-p_i^0)(1-p_j^0)}\Big)\Bigg]\end{aligned}$$

10.10 ESTIMATED VARIANCE (IPW)

$$\begin{aligned}\mathbb{V}[X_T] = \frac{1}{n^2}\Bigg[\sum_{i=1}^{N}\Big(\frac{A_i(Y_i)^2}{(p_i^0)^2} + \frac{(1-A_i)(Y_i)^2}{(1-p_i^0)^2}\Big) - n\hat{\tau}^2 \\ + 2\sum_{i,j\in S}\rho_{ij}\cdot\Big(\frac{A_iA_jY_iY_j}{(p_i^0)^2(p_j^0)^2} + \frac{A_i(1-A_j)Y_iY_j}{(p_i^0)^2(1-p_j^0)^2} \\ + \frac{(1-A_i)A_jY_iY_j}{(1-p_i^0)^2(p_j^0)^2} + \frac{(1-A_i)(1-A_j)Y_iY_j}{(1-p_i^0)^2(1-p_j^0)^2}\Big)\Bigg]\end{aligned}$$

10.11 P IS A RV - FUNCTION OF COVARIATES V

10.11.1 Variance

$\rho(V_i) = P(A=1|V_i)$

$$\begin{aligned}\mathbb{V}[X_T|V] = \frac{1}{n^2}\Bigg[\sum_{i=1}^{N}\Big(\frac{\mathbb{E}[Y_i(1)^2|V_i]}{\rho(V_i)} + \frac{\mathbb{E}[Y_i(0)^2|V_i]}{1-\rho(V_i)}\Big) - n\mathbb{E}[Y(1)-Y(0)|V_i]^2 \\ + 2\sum_{i,j\in S}\rho_{ij}\cdot\Big(\frac{\mathbb{E}[Y_i(1)|V_i]\cdot\mathbb{E}[Y_j(1)]|V_j]}{\rho(V_i)\rho(V_j)} + \frac{\mathbb{E}[Y_i(1)|V_i]\cdot\mathbb{E}[Y_j(0)|V_j]]}{\rho(V_i)(1-\rho(V_j))} \\ + \frac{\mathbb{E}[Y_i(0|V_i)]\cdot\mathbb{E}[Y_j(1)|V_j]]}{(1-\rho(V_i))\rho(V_j)} + \frac{\mathbb{E}[Y_i(0|V_i)]\cdot\mathbb{E}[Y_j(0)|V_j]}{(1-\rho(V_i))(1-\rho(V_j)}\Big)\Bigg]\end{aligned}$$

10.11.2 Estimated Variance (IPW)

$$\begin{aligned}\mathbb{V}[X_T \mid V] = \frac{1}{n^2}\Bigg[\sum_{i=1}^{N}\Big(\frac{A_i\,(Y_i)^2}{\big(\rho(V_i)\big)^2} + \frac{\big(1-A_i\big)\,(Y_i)^2}{\big(1-\rho(V_i)\big)^2}\Big) - n\,\hat{\tau}(V)^2 \\ + 2\sum_{i,j\in S}\rho_{ij}(V_i,V_j)\times\Big(\frac{A_i\,A_j\,Y_i\,Y_j}{\big(\rho(V_i)\big)^2\big(\rho(V_j)\big)^2} + \frac{A_i\,(1-A_j)\,Y_i\,Y_j}{\big(\rho(V_i)\big)^2\big(1-\rho(V_j)\big)^2} \\ + \frac{\big(1-A_i\big)\,A_j\,Y_i\,Y_j}{\big(1-\rho(V_i)\big)^2\big(\rho(V_j)\big)^2} + \frac{\big(1-A_i\big)\,\big(1-A_j\big)\,Y_i\,Y_j}{\big(1-\rho(V_i)\big)^2\big(1-\rho(V_j)\big)^2}\Big)\Bigg].\end{aligned}$$

11 ASYMPTOTIC NORMALITY OF ESTIMATOR

Define

$$x_t \;=\; \frac{1}{n}\sum_{i=1}^{n}\left(\frac{p_i^{(t)}}{p_i^{(0)}}\,Y_i(1) \;-\; \frac{1-p_i^{(t)}}{1-p_i^{(0)}}\,Y_i(0)\right),$$

where x_0 is the estimator using the initial fractional assignments and $x_T = \tau_{\text{swap}}$ is the final binary estimator. Since $\mathbb{E}[\,x_t \mid x_{t-1}\,] = x_{t-1}$, the sequence $\{x_t\}_{t=0}^{T}$ forms a martingale. Let

$$\Delta_t \;=\; x_t \;-\; x_{t-1}.$$

ASSUMPTIONS

Assume that:

- **Super-Population Assumption:** The units $\big(Y_i(1),\, Y_i(0),\, V_i\big)$ are i.i.d. draws from a large (hypothetically infinite) population.
- **Bounded Moments:** $Y_i(1)$ and $Y_i(0)$ have uniformly bounded second moments by a constant $M < \infty$.
- **Bounded Probabilities:** There exists $\varepsilon > 0$ such that

$$\varepsilon \;\le\; p_i^{(0)} \;\le\; 1-\varepsilon \quad \text{for all } i.$$

- **Lower Bound on Quadratic Variation:** There exists a constant $c > 0$ such that for all sufficiently large n,

$$\big\langle x \big\rangle_T \;=\; \frac{1}{n}\sum_{t=1}^{T}\mathbb{E}\big[\Delta_t^2 \mid x_{t-1}\big] \;\ge\; c\,.$$

QUADRATIC VARIATION AND LINDEBERG CONDITION

By a law of large numbers for martingale difference arrays under our boundedness assumptions, we have

$$\big\langle x \big\rangle_T \;\xrightarrow{P}\; \sigma^2 \;>\; 0,$$

with σ^2 as specified in Equation (6) of the paper. Moreover, since each Δ_t is bounded (because of bounded potential outcomes combined with bounded probabilities),

$$\frac{1}{n}\sum_{t=1}^{T}\mathbb{E}\big[\Delta_t^2\,\mathbf{1}_{\{\,|\Delta_t|>\varepsilon\sqrt{n}\,\}} \mid x_{t-1}\big] \;\longrightarrow\; 0,$$

so the Lindeberg condition holds.

MARTINGALE CLT APPLICATION

By the Martingale Central Limit Theorem,

$$\sqrt{n}\,\big(x_T - x_0\big) \;\xrightarrow{d}\; \mathcal{N}\big(0,\sigma^2\big).$$

Since x_0 is the estimator using the initial fractional assignments and $x_T = \tau_{\text{swap}}$, it follows that

$$\sqrt{n}\,\big(\tau_{\text{swap}} - \tau\big) \;\xrightarrow{d}\; \mathcal{N}\big(0,\sigma^2\big).$$

12 PROOF OF LINEAR ESTIMATOR UNBIASEDNESS AND VARIANCE REDUCTION

Proof. Since swap rounding is constructed so that the marginal probability of each treatment indicator is preserved, we have

$$\mathbb{E}[A_i] = p_i \quad \text{for all } i.$$

Therefore, by linearity of expectation,

$$\mathbb{E}[\psi_{swap}] = \mathbb{E}\left[\sum_{i=1}^{n} a_i A_i + b\right] = \sum_{i=1}^{n} a_i \, \mathbb{E}[A_i] + b = \sum_{i=1}^{n} a_i \, p_i + b.$$

Next, to analyze the variance of ψ_{swap}, note that since ψ_{swap} is a linear function of the A_i, its variance is given by

$$\mathbb{V}(\psi_{swap}) = \mathbb{V}\left(\sum_{i=1}^{n} a_i A_i\right) = \sum_{i=1}^{n} a_i^2 \, \mathbb{V}(A_i) + 2 \sum_{i<j} a_i \, a_j \, \text{cov}(A_i, A_j).$$

The swap rounding procedure guarantees that for $i \neq j$,

$$\text{cov}(A_i, A_j) \leq 0.$$

Since we are now assuming that each a_i is non-negative, it follows that each product $a_i \, a_j$ is non-negative. Hence, the cross-covariance terms satisfy

$$2 \sum_{i<j} a_i \, a_j \, \text{cov}(A_i, A_j) \leq 0.$$

It immediately follows that

$$\mathbb{V}(\psi_{swap}) \leq \sum_{i=1}^{n} a_i^2 \, \mathbb{V}(A_i).$$

This upper bound is exactly the variance one would obtain if the A_i were independent (since then all covariances vanish). Thus, the variance under swap rounding is no larger than that under independent assignment. □

PROOF OF VARIANCE REDUCTION VIA COVARIATE-ORDERED PAIRING UNDER A BI-LIPSCHITZ ASSUMPTION

For each unit i, define the effective weight

$$M_i \equiv M(V_i) = \frac{f(V_i) + \tau(V_i)}{p(V_i)} + \frac{f(V_i)}{1 - p(V_i)},$$

and for any pair (i, j) let

$$\overline{M}_{ij} = \frac{M_i + M_j}{2}.$$

It is straightforward to verify that

$$M_i M_j = \overline{M}_{ij}^2 - \left(\frac{M_i - M_j}{2}\right)^2. \tag{28}$$

Thus, for a fixed pair-average $\overline{M}_{ij}$, the product $M_i M_j$ is larger when the absolute difference $|M_i - M_j|$ is smaller.

To compare two pairing schemes, we impose the following assumptions:

(A1) Bi-Lipschitz Condition on $M(\cdot)$: Assume that the mapping

$$M : v \mapsto \frac{f(v) + \tau(v)}{p(v)} + \frac{f(v)}{1 - p(v)}$$

is bi-Lipschitz on the support of the covariates. That is, there exist constants $\ell_M > 0$ and $L_M > 0$ such that for all v, v',

$$\ell_M \|v - v'\| \leq |M(v) - M(v')| \leq L_M \|v - v'\|. \tag{29}$$

(A2) Assumptions on Pairing Distances:

(i) In an *optimal (covariate-ordered) pairing* S_{opt}, assume that every pair (i, j) satisfies

$$\|V_i - V_j\| \leq \delta, \tag{30}$$

for some small $\delta > 0$.

(ii) In an *arbitrary (random) pairing* S_{rand}, assume that the expected squared distance between paired covariates is bounded below:

$$\mathbb{E}_{(i,j)\sim S_{\text{rand}}}\left[\|V_i - V_j\|^2\right] \geq c^2, \tag{31}$$

for some constant $c > 0$ (with the natural requirement that $c > \delta$).

Our goal is to show that under these assumptions the differences $|M_i - M_j|$ in the optimal pairing are uniformly smaller than those in an arbitrary pairing. Consequently, for a fixed pair-average, the product $M_i M_j$ (and thus the negative covariance contributions in the variance of the estimator) is more favorable under optimal pairing.

STEP 1: UPPER BOUND UNDER OPTIMAL PAIRING

For any pair $(i, j) \in S_{\text{opt}}$, by the upper Lipschitz bound in (29) and (30) we have

$$|M_i - M_j| \leq L_M \|V_i - V_j\| \leq L_M \delta.$$

Define

$$\Delta_{\text{opt}} \equiv L_M \delta.$$

Thus, for every optimally paired (i, j) it holds that

$$|M_i - M_j| \leq \Delta_{\text{opt}}.$$

STEP 2: LOWER BOUND UNDER ARBITRARY PAIRING

For any pair (i, j) drawn according to the arbitrary pairing S_{rand}, the lower Lipschitz condition in (29) implies

$$(M_i - M_j)^2 \geq \ell_M^2 \, \|V_i - V_j\|^2.$$

Taking expectations over the random pairing, we obtain

$$\Delta_{\text{rand}}^2 \equiv \mathbb{E}_{(i,j)\sim S_{\text{rand}}}\Big[(M_i - M_j)^2\Big] \geq \ell_M^2 \, \mathbb{E}_{(i,j)\sim S_{\text{rand}}}\Big[\|V_i - V_j\|^2\Big] \geq \ell_M^2 \, c^2.$$

Define

$$\Delta_{\text{rand}} \equiv \sqrt{\Delta_{\text{rand}}^2} \geq \ell_M \, c.$$

STEP 3: COMPARISON OF Δ_{opt} AND Δ_{rand}

If

$$L_M \, \delta < \ell_M \, c, \tag{32}$$

then it follows that

$$\Delta_{\text{opt}} < \Delta_{\text{rand}}.$$

In other words, the differences $|M_i - M_j|$ in the optimal pairing are uniformly smaller than the expected differences under an arbitrary pairing.

STEP 4: IMPLICATION FOR THE PRODUCT $M_i M_j$

Recall the identity in (28):

$$M_i M_j = \overline{M}_{ij}^2 - \left(\frac{M_i - M_j}{2}\right)^2.$$

For a fixed pair-average $\overline{M}_{ij}$, the product $M_i M_j$ is maximized when $|M_i - M_j|$ is minimized. Hence:

- Under the optimal pairing, for every pair (i, j),

$$M_i M_j \geq \overline{M}_{ij}^2 - \frac{1}{4}\Delta_{\text{opt}}^2.$$

- Under arbitrary pairing, the average product satisfies

$$\mathbb{E}_{S_{\text{rand}}}[M_i M_j] \leq \mathbb{E}\Big[\overline{M}_{ij}^2\Big] - \frac{1}{4}\Delta_{\text{rand}}^2.$$

Since $\Delta_{\text{rand}}^2 > \Delta_{\text{opt}}^2$ by (32), the optimal pairing yields, on average, a higher value of $M_i M_j$.

STEP 5: CONNECTION TO VARIANCE REDUCTION

In our setting, the covariance contribution to the variance of an inverse propensity weighted (IPW) estimator is proportional to

$$\text{Cov}(X_i, X_j) = \rho_{ij} \, M_i \, M_j,$$

where $\rho_{ij} < 0$ is induced by the swap rounding procedure. A larger product $M_i M_j$ (in absolute value) means that the negative covariance term is more strongly negative. In turn, this leads to a reduction in the overall variance of the estimator.

CONCLUSION

Under the modified assumptions—namely, that the effective weight function $M(\cdot)$ is bi-Lipschitz (with constants ℓ_M and L_M) and that arbitrary pairings satisfy

$$\mathbb{E}\Big[\|V_i - V_j\|^2\Big] \geq c^2,$$

while the optimal pairing enforces $\|V_i - V_j\| \leq \delta$—we obtain

$$\Delta_{\text{opt}} = L_M\,\delta < \ell_M\,c \leq \Delta_{\text{rand}}.$$

Consequently, the product $M_i M_j$ is higher on average under the optimal (covariate-ordered) pairing than under an arbitrary pairing. This implies that the negative covariance term $\rho_{ij}\,M_i M_j$ is more pronounced, leading directly to a lower overall variance of the IPW estimator.