Online Balanced Experimental Design

David Arbour [*1] **Drew Dimmery** [*2] **Tung Mai** [1] **Anup Rao** [1]

Abstract

We consider the experimental design problem in an online environment, an important practical task for reducing the variance of estimates in randomized experiments which allows for greater precision, and in turn, improved decision making. In this work, we present algorithms that build on recent advances in online discrepancy minimization which accommodate both arbitrary treatment probabilities and multiple treatments. The proposed algorithms are computationally efficient, minimize covariate imbalance, and include randomization which enables robustness to misspecification. We provide worst case bounds on the expected mean squared error of the causal estimate and show that the proposed estimator is no worse than an implicit ridge regression, which are within a logarithmic factor of the best known results for offline experimental design. We conclude with a detailed simulation study showing favorable results relative to complete randomization as well as to offline methods for experimental design with time complexities exceeding our algorithm, which has a linear dependence on the number of observations, by polynomial factors.

1. Introduction

Randomized experimentation is a fundamental tool for obtaining counterfactual estimates. The efficacy of randomization comes from a very simple intuition–by randomly assigning treatment status dependence between observed (and unobserved) treatment and pre-treatment covariates necessarily tends to zero as a function of the number of units. In the context of experimentation, this independence condition on observed covariates, commonly known as balance (Imai et al., 2008; Imai and Ratkovic, 2014), reduces the variance of estimates of the average treatment effect (Greevy et al., 2004; Higgins et al., 2016; Kallus, 2018; Li et al., 2018; Harshaw et al., 2020). Under the appropriate conditions such corrections can result in large increases in effective sample size, allowing for the detection of the small effects which are commonplace in many large scale studies and industrial applications (Dimmery et al., 2019; Azevedo et al., 2020). These contexts rely heavily on experimentation for decision-making, so reduced variance directly translates into more reliable decisions (Kohavi et al., 2012).

Traditional experimental design like blocking (Greevy et al., 2004; Higgins et al., 2016) or even the novel Gram-Schmidt Walk design (Harshaw et al., 2020) require more than one pass over the sample and their sample complexity is greater than $\mathcal{O}(n)$. Even algorithms which admit sequential assignment such as Moore and Moore (2017) suffer from the fact that the algorithm is not linear time and, thus, respondents late in the experiment may take substantially longer to receive a treatment assignment (Cavaille, 2018). Our work is motivated by this setting to provide a linear-time, single-pass (i.e. sequential) algorithm for balancing experimental design. Our focus is on linear measures of balance (often of particular interest to applied researchers). This provides a new avenue through which experimenters can ensure that their experiments optimize the information they gain from costly samples.

We start by presenting four desiderata for effective, practical online experimental design. First, a method must be computationally efficient. In a review of existing methods for online assignment in the case of survey experiments, Cavaille (2018) finds that existing methods become slow to unusable as increasingly more respondents are included in a study. This resulting speed is fundamentally incompatible with effective administration of an experiment. In short, high latency will cause disproportionately large dropoff in an experiment, which may completely nullify the gains from using more sophisticated experimental design. Any algorithm with greater than linear time complexity exacerbates this problem: higher latency for later subjects than earlier subjects will tend to cause non-random sample attrition, as later subjects (who may be different than earlier respondents) will be more likely to drop out.

Second, an experimental design must reduce covariate im-

[*]Equal contribution [1]Adobe Research, San Jose, CA, USA [2]Data Science @ University of Vienna, Vienna, AT. Correspondence to: David Arbour <darbour26@gmail.com>.

Proceedings of the 39^{th} International Conference on Machine Learning, Baltimore, Maryland, USA, PMLR 162, 2022.

balance to be effective. This is the entire justification for using methods more sophisticated than Bernoulli randomization, so if the design is not able to improve on balance, then there will be no subsequent reduction in variance and therefore no compelling reason to use it.

Third, the design must incorporate randomization. Harshaw et al. (2020) provides an extensive discussion on the inherent tradeoff between robustness and balance within experimental design. A design which solely optimizes for balance will tend to operate on a knife-edge of accidental bias (Efron, 1971): the potential bias from an adversarially chosen confounder. With higher accidental bias than Bernoulli randomization, if units do not arrive precisely i.i.d., then the entire design may be compromised. Given that experimental settings are prized specifically for their unbiasedness, this could completely undermine any gains from improved balance. Strong theoretical guarantees on robustness are extremely important in this setting in order to ensure that inferences rest squarely on the design. If assumptions about sampling procedures or data generating processes are necessary to ensure the reliability of estimation, then inferences are not based solely on properties of the experiment, but rather on factors outside the control of the experimenter (Aronow et al., 2021).

Fourth, units which do not show up in the sample should not be included in the balancing. An offline algorithm used in an online environment would fail this condition, because to use it would require to balance the entire population of units *expected* to show up to an experiment. Units within that population who fail to show would nevertheless be assigned a treatment. Depending on the distribution of units who actually show up in the sample, there are no longer any guarantees that balance will obtain.

Our approach satisfies all four of these desiderata. Our contributions are the following:

- We propose an online method for covariate balancing requiring linear time and space which provably provides variance reduction.
- All analyses incorporate an adversarial, non-i.i.d. sampling mechanism in the analysis of worst-case behavior.
- Building upon recent work on discrepancy minimization—the self-balancing walk (Alweiss et al., 2020) and kernel thinning (Dwivedi and Mackey, 2021)—we provide an algorithm whose L_2-imbalance matches the best known *online* discrepancy algorithm. Where δ is a failure probability, performing this optimization online results in a $\log(n/\delta)$ cost in convergence of the average treatment effect over the offline algorithm of discrepancy minimization by Harshaw et al. (2020).
- Using restarts, we show that the *unconditional* performance of this algorithm differs by only a constant factor.
- We extend this algorithm to multiple treatments and non-uniform treatment probabilities.
- These algorithms provide a tuning parameter, ϕ, which allows practitioners to directly manage the balance-robustness tradeoff.

The rest of the paper is organized as follows. Section 2 with a discussion of related work to position our contribution in the literature. Section 3 defines notation and formally introduces the problem. Section 4 provides our proposed algorithms and methods. Section 6 provides a detailed simulation study of the behavior of the proposed algorithms.

2. Related Work

There are two common approaches for achieving improved covariate balance in experiments. The first, and most common especially within industrial settings, approach is to perform a post-hoc regression adjustment which includes pre-treatment covariates (Deng et al., 2013; Lin, 2013). The second approach is to consider covariate balance during the design phase of the experiment, i.e., explicitly optimizing treatment assignment in order to minimize imbalance between treatment groups (Greevy et al., 2004; Higgins et al., 2016; Kallus, 2018).

Post-hoc stratification may be seen as asymptotically equivalent in terms of variance reduction to its analogous pre-stratified design as shown by Miratrix et al. (2013). Miratrix et al.'s (2013) analysis is limited by two factors: it assumes a fixed number of stratification cells (that do not grow with sample size) and it is conditioned on the post-stratification estimator being defined (e.g. treatment and control units within each stratification cell). These limitations may weaken it's asymptotic equivalence argument. A key limitation of post-hoc adjustment approaches is that the desire for simplicity and scalability implies that practitioners typically adjust for only a linear function of the pre-treatment covariates. Indeed, in the common "CUPED" approach, adjustment is performed solely on a linear function of a single pre-treatment outcome measurement (Deng et al., 2013). Second, many common approaches for constructing stratification cells (i.e. clustering algorithms) may be computationally infeasible in practice for industrial applications when the number of simultaneous experiments and the number of outcome variables of interest are large. Third, without sample splitting (or when naively applied) advanced machine-learning based methods for adjustment may slip in assumptions of correct-specification of the outcome model, or have confidence intervals with poor coverage properties. While cross-fitting may ameliorate some

of these problems in larger samples, sample splitting may prove too high a cost when sample sizes are low.

A key shortcoming of design-based covariate balance is the lack of computationally efficient algorithms which provide theoretical guarantees over worst case behavior. Blocking (Fisher, 1935) partitions variables into non-overlapping sets and performs complete randomization within each partition ("blocks"). Higgins et al. (2016) introduced a computationally approximation of blocking which runs in $\mathcal{O}(n\log(n))$ time. Kallus (2018); Bertsimas et al. (2015) propose an optimization based approach, Kallus (2018) additionally considers a partially random approach using semi-definite programming. Zhou et al. (2018) provide a method combining *batch-based* sequential experimentation with rerandomization to achieve balance, but which is not computationally feasible in moderate to large sample sizes. Perhaps the closest to the current work is Harshaw et al. (2020) which propose a balancing design using the Gram-Schmidt walk, an offline method for (linear) discrepancy minimization. Current state of the art for balancing treatment assignment requires polynomial running time and generally requires knowing all of the covariate vectors prior to determining assignment (Higgins et al., 2016; Harshaw et al., 2020; Arbour et al., 2021). As we discuss in section 3, this is a non-starter for online treatment assignment. In this setting, subjects must be allocated as they arrive; it does no good to know how you *should have* assigned a user at the end of the experiment; you need to know when that subject arrives. It's crucial that when a subject in an experiment arrives they be swiftly allocated to a unit. Especially in an online environment, high latency will lead to attrition, which may counteract any potential gains from greater efficiency. Moreover, the users who attrit may be the very subjects of interest (Munger et al., 2021). By inducing differential attrition based on patience, the sample in the experiment may differ greatly from the population of interest on unobserved characteristics that make it difficult to extrapolate to a population-level effect (Egami and Hartman, 2020).

There is also a variety of methods aimed at sequential, online assignment in experiments. The seminal work in this literature is Efron (1971) which introduced an online variant of complete randomization which aims to ensure that a pre-specified marginal treatment probability is met without introducing too much accidental bias. Smith (1984) provides a generalization of the Efron (1971) approach which extends gracefully to multiple treatments. There are a variety of online balanced coin designs which seek to reduce covariate imbalance (e.g. Baldi Antognini and Zagoraiou, 2011; Moore and Moore, 2017). Moore and Moore (2017) is based around Mahalanobis distance. As such, it has polynomial time-complexity *at each arrival time*. In addition to inefficiency, the theoretical worst-case behavior of this algorithm has not been resolved, even in the stochastic setting. Theoretical guarantees of this sort are paramount in the design setting, as practitioners need to know the credibility of their inferences (and how they may differ from simple Bernoulli randomization).

3. Background and Problem Setting

We first fix notation. Random variables will be denoted in upper case, with sets in bold. The problem setting, which we refer to as experimental treatment allocation, is as follows. We assume that we observe $1,\dots,n$ i.i.d. observations of $\boldsymbol{X} \in \mathbb{R}^{n\times d}$: the covariates[1]. The experimenter is asked to assign a treatment assignment, $A \in \{1,-1\}$ (we will later loosen this to multiple discrete treatment values). We will refer to the assignments of A as treatment and control, respectively. Each unit is imbued with *potential outcomes* for each treatment, the value of the outcome if that unit had been assigned to the given group: $y(1)$ for treatment and $y(-1)$ for control. After assignment we observe only the potential outcome corresponding to the realized treatment assignment, Y. We assume that the outcomes are not available until the conclusion of the experiment. At the end of the experiment we are interested in measuring the sample average treatment effect (SATE) between any two treatments, k and k' with the difference in means estimator:

$$\hat{\tau}_{kk'} = \frac{1}{n}\sum_{i}^{n}\frac{A_i}{p(A_i)}Y_i \tag{1}$$

where $p(A_i)$ denotes the probability of assigning treatment A_i to instance i. Note that this is simply the difference-in-means rather than the more general Horwitz-Thompson estimator (Horvitz and Thompson, 1952), as the treatment probability is marginal rather than conditional. More sophisticated estimators are usable in this setting (e.g. Tsiatis et al., 2008; Aronow and Middleton, 2013), but we will focus on the simplest as we optimize design as is commonplace for studying design (Kallus, 2017; Harshaw et al., 2020).

If propensity scores are constant, the estimator of the SATE given by equation 1 will be unbiased and consistent for its oracle counterpart,

$$\tau_{kk'} = \frac{1}{n}\sum_{i}^{n} y_i(k) - y_i(k'), \tag{2}$$

the difference of potential outcomes of the k and k' treatments. This SATE is our estimand of interest, as estimated by equation 1.

[1] We assume linear feature maps throughout. We note that nonlinearities can be handled with the same guarantees following Dwivedi and Mackey (2021) at the cost of additional computational complexity.

We will maintain the following assumptions:

Assumption 1 (Consistency). $Y_i = y_i(k)$ *if* $A_i = k \quad \forall i, k.$

The problem of experimental allocation is to observe covariate vectors and assign A to units so as to achieve desirable properties of the SATE (for instance, to minimize variance). In the most general setting where no assumptions are placed the relationship between the covariates and outcome complete randomization—randomly drawing assignments without respect to background covariates—is known to be minimax optimal (Kallus, 2017).

3.1. Robustness in sequential design

Experiments are prized for their ability to provide unbiased estimates of causal effects with relatively mild assumptions. These assumptions are on the *design* of the experiment rather than more difficult assumptions about the data used in the course of analysis (Sekhon, 2009; Aronow et al., 2021).

In the study of vector balancing, there are three main sampling schemes of interest, listed in order of how adversarial they are:

- *Stochastic arrivals* Units are sampled i.i.d. from some fixed (possibly infinite) population. As such, a given covariate vector is just as likely to arrive early in the sequence as late.
- *Oblivious adversarial arrivals* The adversary is allowed to arbitrarily set the order in which units arrive prior to the first assignment, but cannot change the order in response to assignment of the units to treatment/control.
- *Fully adversarial arrivals* The adversary is allowed to arbitrarily set the order in which units arrive and may change the ordering in response to assignments. The fully adversarial case has a discrepancy bound of $\Omega\left(\sqrt{T}\right)$, with Bernoulli randomization as the optimal strategy (Alweiss et al., 2020).

Assumption 2 (Oblivious Adversary). *The current imbalance is independent of the newly arrived covariate profile conditional on the history of covariate profiles*

3.2. Balance v. Robustness

The fully adversarial case represents an extreme form of robustness. In the face of such an adversary there is no strategy more effective than complete randomization, as in the No Free Lunch Theorem of Kallus (2018). As such, it isn't possible to reduce imbalance by departing from complete randomization without incurring a large cost in terms of robustness. In the oblivious adversarial setting, however, the best available algorithms for discrepancy minimization incorporate randomization (Alweiss et al., 2020; Dwivedi and Mackey, 2021), but with substantially less entropy than in the fully adversarial environment by modulating their assignment probabilities based on the current state of imbalance, leading to $\mathcal{O}(\log nd/\delta)$ discrepancy with probability $1-\delta$. In this setting, it is not necessary to "purely" randomize to ensure sufficient robustness, but rather it is only necessary to incorporate *sufficient* randomization so that an adversary cannot recover too much information about the state of the system at any point in time.

It is this setting which incurs a substantive tradeoff between imbalance minimization and robustness (Harshaw et al., 2020): too much randomization will result in greater imbalance (and, therefore, variance), while too little will result in the potential for exploitation by the adversary. The algorithmic task is to assign units such that imbalance is reduced, while randomizing enough that it's impossible to infer the algorithm's state at any time. Our approach incorporates a parameter, ϕ, to manage the tradeoff between imbalance reduction and worst-case error. This allows the practitioner to choose how to manage this tradeoff based on their preferences between the two.

4. Weighted Online Discrepancy Minimization

Our approach for online experimental design uses weighted online discrepancy minimization. In what follows, we first describe our procedure–a variant of prior work (Alweiss et al., 2020; Dwivedi and Mackey, 2021) to accommodate arbitrary marginal treatment probabilities–before describing it's application and implications for experimental design.

We consider a variant of the online Komlós problem (Spencer, 1977), where vectors $\boldsymbol{x}_1, \dots \boldsymbol{x}_n$ arrive one by one and must be immediately assigned a weighted sign of either $-2q$ or $2(1-q)$, for $0 < q < 1$, such that the weighed discrepancy $\|\sum_{i=1}^{n} \eta_i \boldsymbol{x}_i\|_\infty$, where η_i is the weighted sign given to $\boldsymbol{x}_i$, is minimized. Notice that when $q = 1/2$, the signs become ± 1.

Algorithm 1, takes $i = 1, \dots, n$ unit vectors in sequentially and assigns them to a treatment and control, represented by the value of η_i. The procedure, an extension of recent work in online discrepancy minimization (Alweiss et al., 2020; Dwivedi and Mackey, 2021), assigns treatment with probability proportional to the inner product between a running sum of the signed prior observations. The algorithm and analysis differs from prior work for discrepancy in two aspects which are necessary for use in experimentation:

1. We provide a ridge regression guarantee, by characteriz-

ing the random vectors output by the algorithm in terms of the projection matrix, $\boldsymbol{P}_i = \boldsymbol{X}_i^\top \left(\boldsymbol{X}_i \boldsymbol{X}_i^\top\right)^{-1} \boldsymbol{X}_i$, where $\boldsymbol{X}_i$ is the $i \times d$ submatrix of $\boldsymbol{X}$ corresponding to covariates $\{\boldsymbol{x}_1, .., \boldsymbol{x}_i\}$. This can be seen as an online analogue to what is provided Harshaw et al. (2020) for offline discrepancy minimization.

2. A generalization of the discrepancy minimization algorithm given by Dwivedi and Mackey (2021) to allow for arbitary marginal probabilities q. This allows for the case of imbalanced treatment assignments. A straightforward adoption of the analysis in Dwivedi and Mackey (2021) to this case results in a worse dependence on $1/q$. We derive a sub-exponential concentration bound and get a better dependence on $1/q$.

Algorithm 1 takes each input vector $\boldsymbol{x}_i$ and assigns it $\{-2q, 2(1-q)\}$ signs online to maintain low weighted discrepancy with probability $1-\delta$.

Input: $\boldsymbol{x}, q$
$c \leftarrow \min(1/q, 9.3)\log(2n/\delta)$
for i from 1 to n **do**
  if $|\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i| > c$ **then**
    $\boldsymbol{w}_i \leftarrow \boldsymbol{w}_{i-1} - 2q\frac{\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i}{c}\boldsymbol{x}_i$
  else
    $\eta_i \leftarrow \begin{cases} 2(1-q), & \text{w.p. } q(1-\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i/c) \\ -2q, & \text{w.p. } 1-q(1-\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i/c) \end{cases}$
    $\boldsymbol{w}_i \leftarrow \boldsymbol{w}_{i-1} + \eta_i \boldsymbol{x}_i$
  end if
end for
Output: $\boldsymbol{\eta}, \boldsymbol{w}$

Our main theorem relies on two non-standard definitions of sub-Gaussian and sub-exponential random vectors. We provide a definition of each below before introducing our main results.

Definition 1 (Sub-Gaussian)**.** *A mean zero random vector $\boldsymbol{w}$ is $(\sigma, \boldsymbol{P})$ sub-Gaussian if for all unit vectors $\boldsymbol{u}$ and $\lambda \in \mathbb{R}$, $\mathrm{E}[\exp\left(\lambda \boldsymbol{w}^\top \boldsymbol{u}\right)] \leq \exp\left(\frac{\lambda^2 \sigma^2 \boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}}{2}\right)$.*

Definition 2 (Sub-exponential)**.** *A mean zero random vector $\boldsymbol{w}$ is $(\nu, \alpha, \boldsymbol{P})$ sub-exponential if for all unit vectors $\boldsymbol{u}$ and $|\lambda| \leq \frac{1}{\alpha\sqrt{\boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}}}$, $\mathrm{E}[\exp\left(\lambda \boldsymbol{w}^\top \boldsymbol{u}\right)] \leq \exp\left(\frac{\lambda^2 \nu^2 \boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}}{2}\right)$.*

Theorem 1 (Main)**.** *Let $\boldsymbol{w}_1, \ldots \boldsymbol{w}_n$ be as in Algorithm 1, $A = 0.5803$, $B = 0.4310$ and $\alpha = 2/B$. Then*

1. *$\boldsymbol{w}_i$ is mean zero $\left(\sqrt{c/2q}, P_i\right)$ sub-Gaussian.*
2. *$\boldsymbol{w}_i$ is mean zero $\left(\sqrt{8Ac}, \alpha, P_i\right)$ sub-exponential.*
3. *With probability $1-\delta$, for all i, $|\boldsymbol{w}_i^T \boldsymbol{x}_i| \leq c$.*

Note that η_i is defined only when $|\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i| \leq c$. Therefore, $\boldsymbol{\eta}$ is defined with probability at least $1-\delta$.

Remark 1. *If $\boldsymbol{w}_i$ is a is a mean zero $(\sigma, \boldsymbol{P}_i)$ sub-Gaussian random vector, then $\mathrm{Cov}(\boldsymbol{w}_i) \leq \sigma^2 \boldsymbol{P}_i$. Similarly, we have that if $\boldsymbol{w}_i$ is a is a mean zero (ν, α, P_i) sub-exponential random vector, then $\mathrm{Cov}(\boldsymbol{w}_i) \leq \frac{3}{2}\nu^2 \boldsymbol{P}_i$. A proof is given in Lemma 6.*

We will use Theorem 1 to derive results on the average treatment effect using the framework developed by (Harshaw et al., 2020).

5. Online Experimental Design

We now describe how the results developed in section 4 can be applied to experimental design, and describe the implications in terms of the efficacy of the estimate.

Let $z_i = \eta_i + 2q - 1 \in \{-1, 1\}$, denote treatment status with 1 indicating treatment, and -1 indicating control, respectively. We will show that the expected imbalance is zero, i.e., $\mathrm{E}[\eta_i] = 0$, and so we have marginal treatment probability $\mathrm{E}[z_i] = 2q - 1$, and $\boldsymbol{z} - \mathrm{E}[\boldsymbol{z}] = \boldsymbol{\eta}$. Let $\boldsymbol{\mu} = \frac{\boldsymbol{Y}(1)}{4q} + \frac{\boldsymbol{Y}(0)}{4(1-q)}$. Harshaw et al. (2020) give a linear algebraic expression for the error of HT-estimators in terms of $\boldsymbol{\mu}$. In particular, they show

Lemma 1 (Lemma A2 and Corollary A1 in (Harshaw et al., 2020))**.** *Let $\hat{\tau} = \frac{1}{n}\sum_i^n \frac{A_i}{p(A_i)} Y_i$ be the empirical estimate of the average treatment effect and τ be its population counterpart. We have*

$$\hat{\tau} - \tau = \frac{2}{n}\left(\boldsymbol{z} - \mathrm{E}[\boldsymbol{z}]\right)^\top \boldsymbol{\mu} = \frac{2}{n}\boldsymbol{\eta}^\top \boldsymbol{\mu}$$

and hence,

$$\mathrm{Var}(\hat{\tau}) = \mathrm{E}\left[(\hat{\tau} - \tau)^2\right] = \frac{4}{n^2}\boldsymbol{\mu}^\top \mathrm{Cov}(\boldsymbol{z})\boldsymbol{\mu}.$$

5.1. Controlling Balance/Robustness Tradeoff

Theorem 1 immediately implies that assignments generated by Algorithm 1 are well balanced. But optimizing just for balance can lead to accidental bias (Efron, 1971). We have from Lemma 1 that $\mathrm{Var}(\hat{\tau}) = \frac{4}{n^2}\lambda_{\max}\left(\mathrm{Cov}(\boldsymbol{z})\right)\|\boldsymbol{\mu}\|^2$ in the worst case when $\boldsymbol{\mu}$ is along the top eigenvector of $\mathrm{Cov}(\boldsymbol{z})$. Therefore, to control accidental bias we need to make sure $\lambda_{\max}\left(\mathrm{Cov}(\boldsymbol{z})\right)$ is not high.

We achieve this by augmenting the original covariates $\boldsymbol{x}_i$ by $\sqrt{\phi}\boldsymbol{e}_i$ to get $\begin{bmatrix} \sqrt{\phi}\boldsymbol{e}_i \\ \sqrt{1-\phi}\boldsymbol{x}_i \end{bmatrix}$, where $\phi \in (0, 1)$ is a parameter which controls the extent of the covariate balance, and $\boldsymbol{e}_i$ is a basis vector in dimension n. By a simple calculation, we can see that running Algorithm 1 on $\boldsymbol{x}_i$ augmented with $\sqrt{\phi}\boldsymbol{e}_i$ is equivalent to running it with $\boldsymbol{x}_i$ and

replacing $\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i$ by $\sqrt{1-\phi}\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i$ everywhere. Therefore, we don't have to explicitly augment the covariates in the algorithm.

We note that with augmented covariates, $\boldsymbol{w}_n = \begin{bmatrix} \sqrt{\phi}\boldsymbol{\eta} \\ \sqrt{1-\phi}\boldsymbol{X}^\top\boldsymbol{\eta} \end{bmatrix}$ is a sub-Gaussian or a sub-exponential random vector as in Theorem 1.

The parameter ϕ controls the worst case behavior of the assignment process. In the fully general case, when no assumptions are placed on the potential outcomes, it's known that complete randomization is minimax optimal. However, as our results and others show, with additional assumptions substantial variance reduction can be achieved. The largest eigenvalue of the covariance matrix of treatments quantifies the extent to which bias can occur from this misspecification. The function of ϕ in our setting (and also in Harshaw et al. (2020)), is to balance between reducing this worst case bias and improving the precision of effect estimates by minimizing imbalance. We will see in Proposition 6, the *worst case* bias becomes unbounded as the algorithm focuses solely on balance, while the improvement from balancing given in Proposition 7 has the opposite dependence on ϕ.

We now describe the properties of treatment effect estimates generated using the weighted online design. Our main guarantees are within the context of an implicit ridge regression, and so we will define $\boldsymbol{Q} = \left(\phi\boldsymbol{I} + (1-\phi)\boldsymbol{X}\boldsymbol{X}^\top\right)^{-1}$.

Proposition 1. *When Algorithm 1 is run with augmented covariates as described above, then $\boldsymbol{\eta} = \boldsymbol{z} - \mathrm{E}[\boldsymbol{z}]$ is a mean zero $(\sqrt{c/2q}, \boldsymbol{Q})$ sub-Gaussian random vector and also, $\boldsymbol{\eta}$ is a mean zero $(\sqrt{8Ac}, \alpha, \boldsymbol{Q})$ sub-exponential random vector. Where $A = 0.5803$ as in theorem 1.*

The consequence of proposition 1 is that $\mathrm{E}[\boldsymbol{\eta}] = 0$ for all i, as given below.

Corollary 1 (Unbiasedness)**.** *When Algorithm 1 is run with augmented covariates, we have with probability at least $1-\delta$, $\mathrm{E}\left[\sum_i \boldsymbol{x}_i\eta_i\right] = 0$.* [2]

We now show how these bounds can be applied to control the balance/robustness trade-off by bounding covariance $\mathrm{Cov}(\boldsymbol{z})$ in Loewner order. We will also let $\sigma^2 = c/2q$ if $c = \log(2n/\delta)/q$ and $\sigma^2 = 12Ac$ if $c = 9.3\log(2n/\delta)$. Where $A = 0.5803$ as in theorem 1, for the remainder of the paper.

Corollary 2 (Eigenvalues of Treatment Covariance)**.** *With probability at least $1-\delta$, Algorithm 1 produces $\boldsymbol{\eta}$ satisfying $\mathrm{Cov}(\boldsymbol{z}) = \mathrm{Cov}(\boldsymbol{\eta}) \preceq \sigma^2\boldsymbol{Q}$.*

[2] Here and in other places, $1-\delta$ is a lower bound on the probability that all η_is are defined in Algorithm 1. The expectations are conditional on this event.

Proposition 2 gives a bound on covariate discrepancy at the end of the experiment.

Proposition 2. *Let $\boldsymbol{w} = \sum_i \eta_i\boldsymbol{x}_i$. With probability at least $1-\delta$,*

$$\|\boldsymbol{w}\|_2 \le \sqrt{d}\,\|\boldsymbol{w}\|_\infty \le \min\left(\frac{1}{q}, 9.3\right)\sqrt{\frac{d\log(4d/\delta)\log(4n/\delta)}{2(1-\phi)\phi}}.$$

5.2. Error bounds

Our final set of results provide bounds on the concentration and worst case mean squared error of the treatment effect estimation. As we note above, we do so by bounding our estimate by an implicit ridge regression.

Proposition 3 (Concentration of ATE)**.** *Algorithm 1 when run with augmented covariates, generates a random assignment $\boldsymbol{z}$ such that*

$$|\hat{\tau} - \tau| = \frac{2}{n}|\boldsymbol{\eta}^\top\boldsymbol{\mu}| \le \frac{2c}{n}\sqrt{\boldsymbol{\mu}^\top\boldsymbol{Q}\boldsymbol{\mu}}.$$

with probability $1-\delta$.

Lemma A10 in (Harshaw et al., 2020) shows that $\boldsymbol{\mu}^\top\boldsymbol{Q}\boldsymbol{\mu} = \min_{\boldsymbol{\beta}\in\mathbb{R}^d}\left[\frac{1}{\phi}\|\boldsymbol{\mu} - \boldsymbol{X}\boldsymbol{\beta}\|^2 + \frac{\|\boldsymbol{\beta}\|^2}{(1-\phi)}\right]$.

Proposition 4. *The worst-case mean squared error of the online balancing walk design is upper bounded by $\mathrm{E}\left[(\hat{\tau}-\tau)^2\right] \le \frac{4\sigma^2}{\phi n^2}\sum_{i=1}^n \mu_i^2$ where $\phi \in (0,1)$ with probability $1-\delta$.*

Proof of Proposition 4. This follows from Lemma 1 and Proposition 2. We note that $\boldsymbol{Q} \preceq \frac{\sigma^2}{\phi}\boldsymbol{I}$. □

Proposition 5 (Ridge Connection)**.** *The worst-case mean squared error of the online balancing walk design is upper bounded by an implicit ridge regression estimator with regularization proportional to ϕ. That is, $\mathrm{E}\left[(\hat{\tau}-\tau)^2\right] \le \frac{4\sigma^2 L}{n}$ where $L = \min_{\boldsymbol{\beta}\in\mathbb{R}^d}\left[\frac{1}{\phi n}\|\boldsymbol{\mu} - \boldsymbol{X}\boldsymbol{\beta}\|^2 + \frac{\|\boldsymbol{\beta}\|^2}{(1-\phi)n}\right]$ with probability $1-\delta$.*

We note that proposition 5 closely mirrors the results obtained by Harshaw et al. (2020) in the offline case, with an additional penalty (σ^2) incurred by using an online procedure.

5.3. Algorithm with Restart

We saw earlier that $\boldsymbol{\eta}, \boldsymbol{z}$ are defined only with probability $1-\delta$. This is because with our choice of c, only with probability $1-\delta$ we have for all $i, |\boldsymbol{w}_i^\top\boldsymbol{x}_i| \le c$. This means our treatment assignment fails with probability δ. There is a simple way to make sure that the algorithm never fails and have same error bound guarantees with a slightly worse constant.

To do so, we slightly modify Algorithm 1 so that whenever $|\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i| > c$ for a particular i, we start a new instance of the algorithm for covariates $\boldsymbol{x}_{i+1}, ...\boldsymbol{x}_n$. This is equivalent to setting $\boldsymbol{w}_{i-1} = 0$ and continuing with the algorithm.

Since for any treat assignment procedure $\mathrm{E}\left(\hat{\tau} - \tau\right)^2$ just depends on $\mathrm{Cov}(\boldsymbol{z})$, and Algorithm 1 fails with probability $\geq \delta$, we can show that

$$\begin{aligned}\mathrm{Cov}(\boldsymbol{z}) &\preceq (1-\delta)\boldsymbol{Q} + \delta\boldsymbol{Q} + \delta^2\boldsymbol{Q} + ... \\ &\preceq 2\boldsymbol{Q} \text{ when } \delta \leq 1/2.\end{aligned}$$

Therefore, for the modified algorithm, we will have error guarantees as in Propositions 4 and 5 (but worse by at most a factor of 2) and with probability one.

5.4. Extension to Multiple Treatments

In this section, we consider an online multi-treatment setting, where each vector is assigned to a group in $M = \{m_1, m_2, \ldots, m_k\}$ immediately on arrival. For each $1 \leq i \leq k$, group m_i is associated with a weight α_i. The goal is to minimize the multi-treatment discrepancy:

$$\max_{m_1, m_2 \in M} 2 \left\| \frac{\boldsymbol{s}(m_1)/\alpha_1 - \boldsymbol{s}(m_2)/\alpha_2}{1/\alpha_1 + 1/\alpha_2} \right\|_\infty$$

where $\boldsymbol{s}(m)$ is the sum of all vectors assigned to treatment m. Notice that by setting $\alpha_1 = \frac{1}{1-q}$ and $\alpha_2 = \frac{1}{q}$, we can recover the definition given for the weighted discrepancy between two treatments m_1 and m_2.

Our algorithm can leverage any oracle (we call it `BinaryBalance` in Algorithm 2) that minimizes the weighted discrepancy for two treatments. Our results are obtained by using Algorithm 1.

We first build a binary tree where each leaf of the tree corresponds to one of the k treatments in M. Let h be the smallest integer such that $2^h \geq k$. We start with a complete binary tree of height h, and then remove $2^h - k$ leaves from the tree such that no two siblings are removed. Note that this is possible by the definition of h. We further contract each internal node with only one child to its child. This process does not change the number of leaves in the tree. Let T be the obtained tree. By construction, all internal nodes of T have 2 children and T has exactly k leaves. We then associate each leaf of T with a treatment in M. For each vector assigned to treatment m_i, we also say that it is assigned to the leaf corresponding to m_i, $\forall 1 \leq i \leq k$. For each node $v \in T$, denote by $\boldsymbol{s}(v)$ the sum of all vectors assigned to leaves under v. In addition, let $\alpha(v)$ be the sum of all weights assigned to leaves under v. For each internal node v of T, the weighted discrepancy vector at v is defined as:

$$\boldsymbol{w}(v) = \frac{\alpha(v_r)}{\alpha(v_l) + \alpha(v_r)} \boldsymbol{s}(v_l) - \frac{\alpha(v_l)}{\alpha(v_l) + \alpha(v_r)} \boldsymbol{s}(v_r),$$

where v_l and v_r are the left and right child of v respectively.

For each internal node v in T, we maintain an independent run of a two-treatment algorithm that minimizes $\|\boldsymbol{w}(v)\|_\infty$. At a high level, we minimize the weighted discrepancies at all internal nodes simultaneously. When a new vector $\boldsymbol{x}$ arrives, we first feed it to the algorithm at root r. If the result is $+$, we continue with the left sub-tree of r. Otherwise, we go to the right sub-tree. We continue in that manner until we reach a leaf l. $\boldsymbol{x}$ will then be assigned to l (and the treatment associated with l).

Theorem 2. *Let* `BinaryBalance` *be Algorithm 1. Then Algorithm 2 obtains* $O\left(\log k \sqrt{\frac{(1-\phi)d\log(dk/\delta)\log(nk/\delta)}{\phi}}\right)$ *multi-treatment discrepancy with probability* $1-\delta$.

Algorithm 2 `KGroupBalance` takes each input vector x_i and assigns it to one of the groups online to maintain low discrepancy with probability $1-\delta$.

Input: $\boldsymbol{x}, k, \alpha$.
$h \leftarrow$ smallest integer such that $2^h \geq k$.
$T \leftarrow$ complete binary tree with height h. Remove $2^h - k$ leaves from T such that no two siblings are removed. Associate each treatment to a leaf of T. Contract each internal node in T with one child to its child.
for node v in T **do**
  $\alpha(v) \leftarrow$ sum of all weights of groups associating with leaves under v.
end for
for internal node v of T **do**
  Instantiate `BinaryBalance`$(v) \leftarrow$ oracle for weighted discrepancy problem at v with weighted signs $\frac{\alpha(v_r)}{\alpha(v_l)+\alpha(v_r)}$ and $-\frac{\alpha(v_l)}{\alpha(v_l)+\alpha(v_r)}$.
end for
for i from 1 to n **do**
  $v \leftarrow$ root of T.
  for v is an internal node of T **do**
    Feed $\boldsymbol{x}_i$ to `BinaryBalance`(v)
    $v \leftarrow$ one of the children of v according to the assignment of `BinaryBalance`(v) on input $\boldsymbol{x}_i$
  end for
  Assign x_i to the group corresponding to v.
end for

6. Experiments

In this section, we provide simulation evidence on the efficacy of our proposed methods. In particular, we use a wide variety of data generating processes, many of which do not assume that units arrive i.i.d. as is often standard in simulation settings for this problem. Subjects are unlikely to

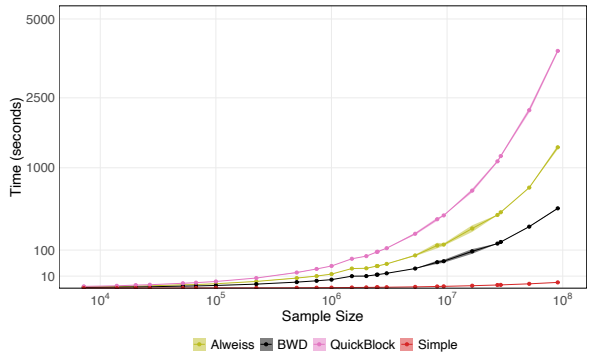

Figure 1. Time to design. All timings performed on a `ml.r5.2xlarge` instance of Amazon SageMaker. The y-axis is scaled by the square root for easier visualization.

arrive truly i.i.d. in the real world. Earlier arrivals will typically be more active than late-arriving units, for example.

All data generating processes used in simulations are shown in Table A1. If not otherwise specified, the sample size is 1000 subjects, the number of groups is two and the marginal probability of treatment is $\frac{1}{2}$.

Methods compared in the simulations are simple randomization (Bernoulli coin flips), complete randomization (fixed-margins randomization), the biased coin designs of Efron (1971) and of Smith (1984), QuickBlock of Higgins et al. (2016), and Alweiss et al. (2020). These are compared to our proposed methods which are generalized versions of the discrepancy minimization procedure of Dwivedi and Mackey (2021). We provide three versions of our proposed algorithms, called BWD for "Balancing Walk Design". The most basic version (BWDRandom) reverts to simple random assignment for all remaining periods once $|\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i| > c$. Our preferred approach restarts the algorithm in this case as described in section 5.3. We examine this with two levels of robustness, ϕ: 0 (purely balancing) and 0.5 (a uniform mix between randomization and balancing): BWD(0) and BWD(0.5) respectively. For further details, see section A.4.

In these comparisons, BWD need not out-perform all methods in all data-generating processes. QuickBlock, for instance, is a fully off-line method, so comparable performance by an online method is noteworthy. In general, Alweiss and BWD will be most effective when the true relationship between covariates and outcome is linear, since they seek linear balance. All plots incorporate 95% confidence intervals.

6.1. Binary treatment

Timing. Our proposed method (BWD) is highly efficient, scaling substantially better than other balancing methods.

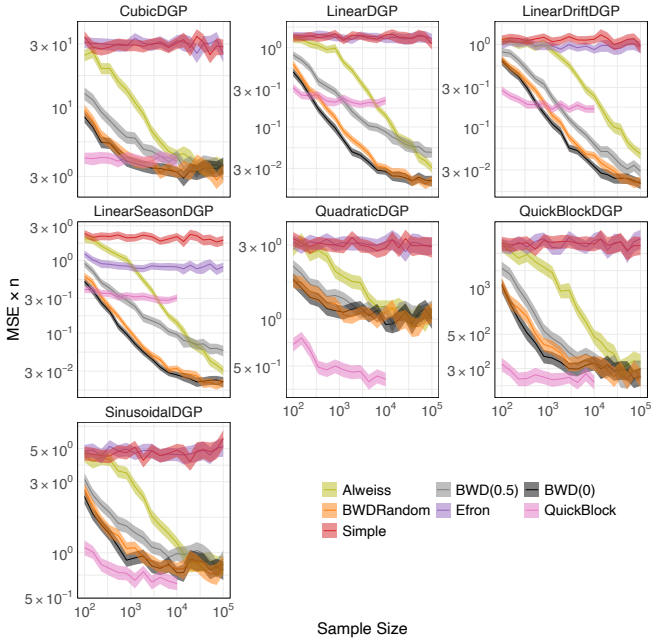

Figure 2. MSE. BWD provides effective variance reduction across a wide array of simulation environments.

This analysis of runtime directly compares online methods to a widely used offline balancing method (QuickBlock). It's important to note that the QuickBlock algorithm cannot be used in the online setting, even if its runtime did not make that prohibitive. While QuickBlock is $\mathcal{O}(n \log n)$, the proposed online balancing methods are all linear-time. Given QuickBlock's runtime, the following simulations only include it for comparisons up to sample sizes of 10^4.

MSE. Next, we demonstrate in Figure 2 how imbalance minimization translates to improved estimation of causal effects, measured by the mean squared-error of our estimate of the average treatment effect. We normalize this graph based on n, the rate of convergence of the difference-in-means estimator under simple randomization. The results depend strongly on the true nature of the data-generating process. In short, on non-linear data generating processes, offline blocking performs better than anything else, but in many settings BWD converges to similar error rates as QuickBlock. On linear or near-linear data-generating processes, our proposed algorithms perform very strongly, outperforming QuickBlock even in small sample-sizes. When there is a break from the purely i.i.d. stochastic setting (such as `LinearDriftDGP` and `LinearSeasonDGP`), BWD behaves well, as expected. When ϕ is chosen to be 0.5 (the line marked `BWD(0.5)`), representing a desire for more emphasis on robustness than on pure imbalance reduction, the reduction in MSE is more

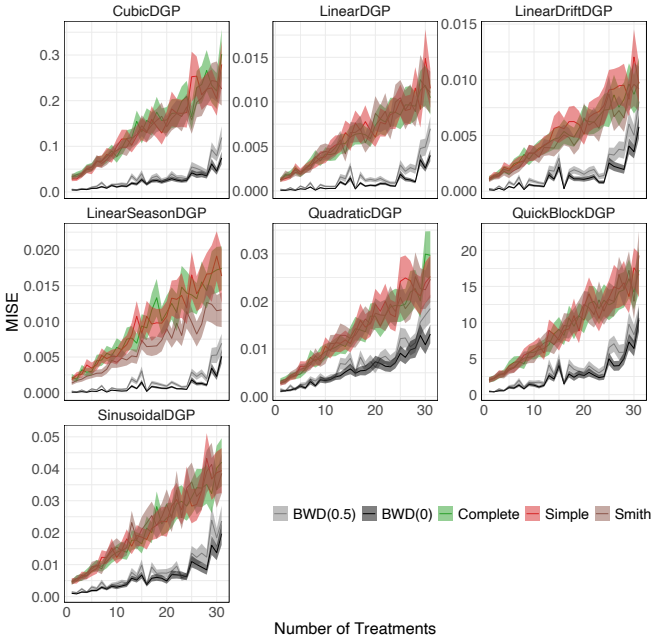

Figure 3. MISE. BWD effective reduces variance no matter the number of treatments.

similar to full imbalance minimization than to complete randomization.

6.2. Multiple Treatments

MISE. BWD gracefully extends to the multiple-treatment setting, which we demonstrate in Figure 3. This chart measures the mean integrated squared-error of our estimates of the ATEs (relative to a single control group). Figure 3 shows the results. BWD consistently outperforms existing online assignment methods by substantial margins.

An array of additional simulation results may be found in Appendix B.

7. Conclusion

Experiments are a crucial part of how humans learn about the world and make decisions. This paper is aimed at providing a way to more effectively run experiments in the online setting. Practitioners must commonly operate their experiments in this environment, but due to the lack of suitable options for design fall back to simple randomization as the assignment mechanism. In this paper, we have shown how the Balancing Walk Design can be an effective tool in this setting. It is efficient, effective at reducing imbalance (and, therefore, the variance of resulting causal estimates), robust and it is fully suited to the particularities of online treatment assignment. Simulations have shown it to work well across a range of diverse settings. The Balancing Walk Design can improve the practice of large-scale online experimentation.

References

R. Alweiss, Y. P. Liu, and M. Sawhney. Discrepancy minimization via a self-balancing walk. *arXiv preprint arXiv:2006.14009*, 2020.

D. Arbour, D. Dimmery, and A. Rao. Efficient balanced treatment assignments for experimentation. In *International Conference on Artificial Intelligence and Statistics*, pages 3070–3078. PMLR, 2021.

P. Aronow, J. M. Robins, T. Saarinen, F. Sävje, and J. Sekhon. Nonparametric identification is not enough, but randomized controlled trials are. *arXiv preprint arXiv:2108.11342*, 2021.

P. M. Aronow and J. A. Middleton. A class of unbiased estimators of the average treatment effect in randomized experiments. *Journal of Causal Inference*, 1(1):135–154, 2013. doi: doi:10.1515/jci-2012-0009. URL https://doi.org/10.1515/jci-2012-0009.

E. M. Azevedo, A. Deng, J. L. Montiel Olea, J. Rao, and E. G. Weyl. A/b testing with fat tails. *Journal of Political Economy*, 128(12):4614–000, 2020.

A. Baldi Antognini and M. Zagoraiou. The covariate-adaptive biased coin design for balancing clinical trials in the presence of prognostic factors. *Biometrika*, 98(3): 519–535, 2011.

D. Bertsimas, M. Johnson, and N. Kallus. The power of optimization over randomization in designing experiments involving small samples. *Operations Research*, 63(4): 868–876, 2015.

C. Cavaille. Implementing blocked randomization in online survey experiments, 2018.

A. Deng, Y. Xu, R. Kohavi, and T. Walker. Improving the sensitivity of online controlled experiments by utilizing pre-experiment data. In *Proceedings of the sixth ACM international conference on Web search and data mining*, pages 123–132, 2013.

D. Dimmery, E. Bakshy, and J. Sekhon. Shrinkage estimators in online experiments. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2914–2922, 2019.

R. Dwivedi and L. Mackey. Kernel thinning. *arXiv preprint arXiv:2105.05842*, 2021.

B. Efron. Forcing a sequential experiment to be balanced. *Biometrika*, 58(3):403–417, 1971. ISSN 00063444. URL http://www.jstor.org/stable/2334377.

N. Egami and E. Hartman. Elements of external validity: Framework, design, and analysis. *SSRN*, 2020.

R. A. Fisher. *The Design of Experiments*. Oliver and Boyd, Edinburgh, 1935.

R. Greevy, B. Lu, J. H. Silber, and P. Rosenbaum. Optimal multivariate matching before randomization. *Biostatistics*, 5(2):263–275, 2004.

C. Harshaw, F. Sävje, D. Spielman, and P. Zhang. Balancing covariates in randomized experiments using the gram-schmidt walk, 2020.

M. J. Higgins, F. Sävje, and J. S. Sekhon. Improving massive experiments with threshold blocking. *Proceedings of the National Academy of Sciences*, 113(27):7369–7376, 2016.

D. G. Horvitz and D. J. Thompson. A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47(260):663–685, 1952. ISSN 01621459. URL http://www.jstor.org/stable/2280784.

K. Imai and M. Ratkovic. Covariate balancing propensity score. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(1):243–263, 2014.

K. Imai, G. King, and E. A. Stuart. Misunderstandings between experimentalists and observationalists about causal inference. *Journal of the Royal Statistical Society. Series A,(Statistics in society)*, 2008.

N. Kallus. Balanced policy evaluation and learning. *arXiv preprint arXiv:1705.07384*, 2017.

N. Kallus. Optimal a priori balance in the design of controlled experiments. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 80(1):85–112, 2018.

R. Kohavi, A. Deng, B. Frasca, R. Longbotham, T. Walker, and Y. Xu. Trustworthy online controlled experiments: Five puzzling outcomes explained. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 786–794. ACM, 2012.

X. Li, P. Ding, and D. B. Rubin. Asymptotic theory of rerandomization in treatment–control experiments. *Proceedings of the National Academy of Sciences*, 115(37):9157–9162, 2018.

W. Lin. Agnostic notes on regression adjustments to experimental data: Reexamining freedman's critique. *The Annals of Applied Statistics*, 7(1):295–318, 2013.

L. W. Miratrix, J. S. Sekhon, and B. Yu. Adjusting treatment effect estimates by post-stratification in randomized experiments. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75(2):369–396, 2013.

R. T. Moore and S. A. Moore. Blocking for sequential political experiments. *Political Analysis*, 21(4):507–523, 2017. doi: 10.1093/pan/mpt007.

K. Munger, I. Gopal, J. Nagler, and J. A. Tucker. Accessibility and generalizability: Are social media effects moderated by age or digital literacy? *Research & Politics*, 8(2):20531680211016968, 2021. doi: 10.1177/20531680211016968. URL https://doi.org/10.1177/20531680211016968.

J. S. Sekhon. Opiates for the matches: Matching methods for causal inference. *Annual Review of Political Science*, 12:487–508, 2009.

R. L. Smith. Sequential treatment allocation using biased coin designs. *Journal of the Royal Statistical Society: Series B (Methodological)*, 46(3):519–543, 1984.

J. Spencer. Balancing games. *Journal of Combinatorial Theory, Series B*, 23(1):68–74, 1977.

A. A. Tsiatis, M. Davidian, M. Zhang, and X. Lu. Covariate adjustment for two-sample treatment comparisons in randomized clinical trials: A principled yet flexible approach. *Statistics in Medicine*, 27(23):4658–4677, 2008. doi: https://doi.org/10.1002/sim.3113. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.3113.

M. J. Wainwright. *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2019. doi: 10.1017/9781108627771.

Q. Zhou, P. A. Ernst, K. L. Morgan, D. B. Rubin, and A. Zhang. Sequential rerandomization. *Biometrika*, 105(3):745–752, 06 2018. ISSN 0006-3444. doi: 10.1093/biomet/asy031. URL https://doi.org/10.1093/biomet/asy031.

A. Proofs

A.1. Proof of Theorem 1

Proposition 6 (Loewner Order)**.** *Let* $\boldsymbol{M} \succeq 0, C \geq 1, L \geq 0$. *If*

$$\boldsymbol{M} \preceq CL\boldsymbol{P} \coloneqq CL\boldsymbol{B}^\top \left(\boldsymbol{B}^\top \boldsymbol{B}\right)^{-1} \boldsymbol{B}$$

then for any vector $\boldsymbol{v} \in \mathbb{R}^n$ *with* $\|\boldsymbol{v}\|_2 \leq 1$

$$\boldsymbol{M}' = \left(\boldsymbol{I} - C^{-1}\boldsymbol{v}\boldsymbol{v}^\top\right) \boldsymbol{M} \left(\boldsymbol{I} - C^{-1}\boldsymbol{v}\boldsymbol{v}^\top\right) + L\boldsymbol{v}\boldsymbol{v}^\top$$

satisfies

$$0 \preceq \boldsymbol{M}' \preceq CL\boldsymbol{P}' \coloneqq CL\boldsymbol{P} + CL\frac{(\boldsymbol{I} - \boldsymbol{P})\,\boldsymbol{v}\boldsymbol{v}^\top\,(\boldsymbol{I} - \boldsymbol{P})}{\boldsymbol{v}^\top\,(\boldsymbol{I} - \boldsymbol{P})\,\boldsymbol{v}}.$$

Proof. By definition of $\boldsymbol{M}'$ and the assumption $\boldsymbol{M} \preceq cL\boldsymbol{P}$, we have

$$\boldsymbol{M}' \preceq CL\left(\boldsymbol{I} - C^{-1}\boldsymbol{v}\boldsymbol{v}^\top\right) \boldsymbol{P} \left(\boldsymbol{I} - C^{-1}\boldsymbol{v}\boldsymbol{v}^\top\right) + L\boldsymbol{v}\boldsymbol{v}^\top.$$

Therefore, it is sufficient to prove

$$\left(\boldsymbol{I} - C^{-1}\boldsymbol{v}\boldsymbol{v}^\top\right) \boldsymbol{P} \left(\boldsymbol{I} - C^{-1}\boldsymbol{v}\boldsymbol{v}^\top\right) + C^{-1}\boldsymbol{v}\boldsymbol{v}^\top \preceq \boldsymbol{P} + \frac{(\boldsymbol{I} - \boldsymbol{P})\,\boldsymbol{v}\boldsymbol{v}^\top\,(\boldsymbol{I} - \boldsymbol{P})}{\boldsymbol{v}^\top\,(\boldsymbol{I} - \boldsymbol{P})\,\boldsymbol{v}}.$$

Also note that, since $\|\boldsymbol{v}\| \leq 1$ and $C \geq 1$, we can absorb C into $\boldsymbol{v}$ (by taking $\boldsymbol{v} \coloneqq \boldsymbol{v}/\sqrt{C}$) and therefore without loss of generality assume that $C = 1$ and $\|\boldsymbol{v}\| \leq 1$. Define:

$$\begin{aligned}
&\boldsymbol{P}_x = \boldsymbol{B}^\top \left(\boldsymbol{B}^\top \boldsymbol{B}\right)^{-1} \boldsymbol{B} && \boldsymbol{Q}_x \coloneqq (\boldsymbol{I} - \boldsymbol{P}) \quad \boldsymbol{P}_v = \boldsymbol{v}\boldsymbol{v}^\top && \boldsymbol{Q}_v = \left(\boldsymbol{I} - \boldsymbol{v}\boldsymbol{v}^\top\right) \\
&\boldsymbol{a} \coloneqq \boldsymbol{P}_x \boldsymbol{v} && \boldsymbol{b} \coloneqq \boldsymbol{Q}_x \boldsymbol{v} \quad \alpha \coloneqq \|\boldsymbol{a}\|^2 && \beta \coloneqq \|\boldsymbol{b}\|^2
\end{aligned}$$

We want to show

$$\boldsymbol{Q}_v \boldsymbol{P}_x \boldsymbol{Q}_v + \boldsymbol{P}_v \preccurlyeq \boldsymbol{P}_x + \frac{\boldsymbol{Q}_x \boldsymbol{P}_v \boldsymbol{Q}_x}{\beta}$$

Beginning by rewriting the LHS

$$\begin{aligned}
&\boldsymbol{Q}_v \boldsymbol{P}_x \boldsymbol{Q}_v + \boldsymbol{P}_v = (\boldsymbol{I} - \boldsymbol{P}_v)\boldsymbol{P}_x(\boldsymbol{I} - \boldsymbol{P}_v) + \boldsymbol{P}_v \\
&= \boldsymbol{P}_x - \boldsymbol{P}_v\boldsymbol{P}_x - \boldsymbol{P}_x\boldsymbol{P}_v + \boldsymbol{P}_v\boldsymbol{P}_x\boldsymbol{P}_v + \boldsymbol{P}_v \\
&= \boldsymbol{P}_x - \boldsymbol{v}\boldsymbol{a}^\top - \boldsymbol{a}\boldsymbol{v}^\top + \boldsymbol{v}\boldsymbol{v}^\top\alpha + \boldsymbol{P}_v \\
&= \boldsymbol{P}_x - (\boldsymbol{a} + \boldsymbol{b})\boldsymbol{a}^\top - \boldsymbol{a}(\boldsymbol{a} + \boldsymbol{b})^\top + (\boldsymbol{a} + \boldsymbol{b})(\boldsymbol{a} + \boldsymbol{b})^\top\alpha + \boldsymbol{P}_v
\end{aligned}$$

Expanding $\boldsymbol{P}_v$ as $\boldsymbol{P}_v = (\boldsymbol{a} + \boldsymbol{b})\,(\boldsymbol{b} + \boldsymbol{a})^\top = \boldsymbol{a}\boldsymbol{a}^\top + \boldsymbol{b}\boldsymbol{b}^\top + \boldsymbol{a}\boldsymbol{b}^\top + \boldsymbol{b}\boldsymbol{a}^\top$ gives

$$\begin{aligned}
&\boldsymbol{P}_x - (\boldsymbol{a} + \boldsymbol{b})\,\boldsymbol{a}^\top - \boldsymbol{a}\,(\boldsymbol{a} + \boldsymbol{b})^\top + (1 + \alpha)\left(\boldsymbol{a}\boldsymbol{a}^\top + \boldsymbol{b}\boldsymbol{b}^\top + \boldsymbol{a}\boldsymbol{b}^\top + \boldsymbol{b}\boldsymbol{a}^\top\right) \\
=&\boldsymbol{P}_x + (\alpha - 1)\boldsymbol{a}\boldsymbol{a}^\top + \alpha\left(\boldsymbol{a}\boldsymbol{b}^\top + \boldsymbol{b}\boldsymbol{a}^\top\right) + (1 + \alpha)\boldsymbol{b}\boldsymbol{b}^\top
\end{aligned}$$

Since $\|\boldsymbol{v}\|^2 \leq 1$, we have $\alpha + \beta \leq 1$. Now considering the difference of the LHS from the RHS after multiplying both sides by β we arrive at

$$\begin{aligned}
\beta(\text{RHS} - \text{LHS}) =&\boldsymbol{b}\boldsymbol{b}^\top(\underbrace{\quad 1 - \beta(1 + \alpha) \quad}_{1 - \beta - \beta\alpha \geq \alpha(1-\beta) \geq \alpha^2}) + \beta(1 - \alpha)\boldsymbol{a}\boldsymbol{a}^\top - \alpha\beta(\boldsymbol{a}\boldsymbol{b}^\top + \boldsymbol{b}\boldsymbol{a}^\top) \\
\succcurlyeq&\alpha^2\boldsymbol{b}\boldsymbol{b}^\top + \beta^2\boldsymbol{a}\boldsymbol{a}^\top - \alpha\beta(\boldsymbol{a}\boldsymbol{b}^\top + \boldsymbol{b}\boldsymbol{a}^\top) \\
=&\,(\alpha\boldsymbol{b} - \beta\boldsymbol{a})\,(\alpha\boldsymbol{b} - \beta\boldsymbol{a})^\top \succcurlyeq 0.
\end{aligned}$$

□

Lemma 2. *When $|\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i| < c$, we have*

$$\mathrm{E}\left[\eta_i\right] = -2q\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i / c.$$

Proof.

$$\begin{aligned}
\mathrm{E}\left[\eta_i\right] &= 2(1-q)\cdot(q(1-\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i/c)) + (-2q)\cdot(1-q(1-\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i/c)) \\
&= 0 + 2(1-q)(-q\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i/c) + (-2q)(q\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i/c) \\
&= -2q\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i/c.
\end{aligned}$$

□

For all i and $\boldsymbol{u} \in \mathbb{R}^d$, we have

$$\begin{aligned}
\langle \boldsymbol{w}_i, \boldsymbol{u}\rangle &= \left\langle \boldsymbol{w}_{i-1}, \boldsymbol{u} - 2q\frac{\boldsymbol{x}_i^\top \boldsymbol{u}}{c}\boldsymbol{x}_i \right\rangle + \epsilon_i \boldsymbol{x}_i^\top \boldsymbol{u} \\
&= \langle \boldsymbol{w}_i, \left(\boldsymbol{I} - \frac{2q}{c}\boldsymbol{x}_i\boldsymbol{x}_i^\top\right)\boldsymbol{u}\rangle + \epsilon_i \boldsymbol{x}_i^\top \boldsymbol{u} \\
&= \langle \boldsymbol{w}_i, \boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}\rangle + \epsilon_i \boldsymbol{x}_i^\top \boldsymbol{u},
\end{aligned}$$

where $\epsilon_i = 0$ if $|\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i| > c$ and $\epsilon_i = \eta_i - E[\eta_i]$ otherwise and $\boldsymbol{Q}_{\boldsymbol{x}_i,c} = \left(\boldsymbol{I} - \frac{2q}{c}\boldsymbol{x}_i\boldsymbol{x}_i^\top\right)$.

Consider the case when $|\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i| \le c$ and $\epsilon_i = \eta_i - E[\eta_i]$. Let $\tilde{q} = q(1-\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i/c)$. Note that $0 \le \tilde{q} \le 2q$. We have

$$\epsilon_i \leftarrow \begin{cases} 2(1-\tilde{q}) & \text{with probability } \tilde{q}, \\ -2\tilde{q} & \text{with probability } 1-\tilde{q}. \end{cases}$$

Definition 3 (Sub-Gaussian)**.** *A mean zero random variable X is sub-Gaussian with parameter σ if for all $\lambda \in \mathbb{R}$,*

$$\mathrm{E}[\exp\left(\lambda X\right)] \le \exp\left(\frac{\lambda^2\sigma^2}{2}\right).$$

A mean zero random vector $\boldsymbol{w}$ is $(\sigma, \boldsymbol{P})$ sub-Gaussian if for all unit vectors $\boldsymbol{u}$ and $\lambda \in \mathbb{R}$,

$$\mathrm{E}[\exp\left(\lambda \boldsymbol{w}^\top \boldsymbol{u}\right)] \le \exp\left(\frac{\lambda^2\sigma^2\boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}}{2}\right).$$

In particular, $\boldsymbol{w}^\top \boldsymbol{u}$ is σ' sub-Gaussian, where $\sigma'^2 = \sigma^2\boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}$.

Definition 4 (Sub-exponential)**.** *A mean zero random variable X is (ν, α) sub-exponential if for all $|\lambda| \le \frac{1}{\alpha}$,*

$$\mathrm{E}[\exp\left(\lambda X\right)] \le \exp\left(\frac{\lambda^2\nu^2}{2}\right).$$

A mean zero random vector $\boldsymbol{w}$ is $(\nu, \alpha, \boldsymbol{P})$ sub-exponential if for all unit vectors $\boldsymbol{u}$ and $|\lambda| \le \frac{1}{\alpha\sqrt{\boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}}}$,

$$\mathrm{E}[\exp\left(\lambda \boldsymbol{w}^\top \boldsymbol{u}\right)] \le \exp\left(\frac{\lambda^2\nu^2\boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}}{2}\right).$$

In particular, $\boldsymbol{w}^\top \boldsymbol{u}$ is (ν', α') sub-exponential, where $\nu'^2 = \nu^2\boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}$ and $\alpha' = \alpha\sqrt{\boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}}$.

The following concentration bounds for sub-Gaussian and sub-exponential vectors are obtained from standard bounds for scalar sub-Gaussian and sub-exponential random variables (Wainwright, 2019) by scaling σ, ν and α by appropriate factors.

Lemma 3 (Sub-Gaussian Concentration)**.** *If a random vector $\boldsymbol{w}$ is $(\sigma, \boldsymbol{P})$ sub-Gaussian, then for all unit vectors $\boldsymbol{u}$,*

$$P\left(|\boldsymbol{w}^\top \boldsymbol{u}| \geq t\right) \leq 2\exp\left(-\frac{t^2}{2\sigma^2 \boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}}\right).$$

Lemma 4 (Sub-exponential Concentration)**.** *If a random vector $\boldsymbol{w}$ is $(\nu, \alpha, \boldsymbol{P})$ sub-exponential, then for all unit vectors $\boldsymbol{u}$,*

$$P\left(|\boldsymbol{w}^\top \boldsymbol{u}| \geq t\right) \leq \begin{cases} 2\exp\left(-\frac{t^2}{2\nu^2 \boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}}\right) & \text{if } 0 \leq t \leq \frac{\nu^2 \boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}}{\alpha} \\ 2\exp\left(-\frac{t}{2\alpha \boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}}\right) & \text{if } t > \frac{\nu^2 \boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}}{\alpha}. \end{cases}$$

Lemma 5. *Suppose X is a (ν, α) sub-exponential random variable with $\frac{2}{\nu^2} \leq \frac{1}{\alpha^2}$. Then*

$$\mathrm{Var}(X) \leq \frac{3}{2}\nu^2.$$

Proof. We will use the inequality $x^2 \leq C\left(e^x + e^{-x}\right), \forall x$ and $C = 1.5/e$. This gives

$$\begin{aligned} \mathrm{Var}(\lambda X) &\leq C\mathrm{E}\left(e^{\lambda X} + e^{-\lambda X}\right) \\ &\leq 2C\exp(0.5\lambda^2\nu^2), \text{ if } |\lambda| \leq \frac{1}{\alpha}. \end{aligned}$$

We therefore have

$$\begin{aligned} \mathrm{Var}(X) &\leq C\nu^2 \frac{\exp(0.5\lambda^2\nu^2)}{0.5\lambda^2\nu^2} \\ &= eC\nu^2 \text{ when } 0.5\lambda^2\nu^2 = 1. \end{aligned}$$

We get the result with $C = 1.5/e$. □

Lemma 6. *If $\boldsymbol{w}$ is a is a mean zero $(\sigma, \boldsymbol{P})$ sub-Gaussian random vector, then $\mathrm{Cov}(\boldsymbol{w}) \preceq \sigma^2 \boldsymbol{P}$. If $\boldsymbol{w}$ is a is a mean zero $(\nu, \alpha, \boldsymbol{P})$ sub-exponential random vector with $\frac{2}{\nu^2} \leq \frac{1}{\alpha^2}$, $\mathrm{Cov}(\boldsymbol{w}) \preceq \frac{3}{2}\nu^2 \boldsymbol{P}$.*

Proof. For all unit vector $\boldsymbol{u}$, we have

$$\boldsymbol{u}^\top \mathrm{Cov}(\boldsymbol{w})\boldsymbol{u} = \boldsymbol{u}^\top \mathrm{E}(\boldsymbol{w}\boldsymbol{w}^\top)\boldsymbol{u} = \mathrm{E}[(\boldsymbol{w}^\top \boldsymbol{u})^2] = \mathrm{Var}(\boldsymbol{w}^\top \boldsymbol{u}).$$

By definition, if $\boldsymbol{w}$ is a is a mean zero $(\sigma, \boldsymbol{P})$ sub-Gaussian random vector, $\boldsymbol{w}^\top \boldsymbol{u}$ is $\sigma^2 \boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}$ sub-Gaussian. Therefore,

$$\boldsymbol{u}^\top \mathrm{Cov}(\boldsymbol{w})\boldsymbol{u} = \mathrm{Var}(\boldsymbol{w}^\top \boldsymbol{u}) \leq \sigma^2 \boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}$$

as desired.

Similarly, if $\boldsymbol{w}$ is a is a mean zero $(\nu, \alpha, \boldsymbol{P})$ sub-exponential random vector, $\boldsymbol{w}^\top \boldsymbol{u}$ is $(\nu\sqrt{\boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}}, \alpha\sqrt{\boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}})$ sub-exponential. By Lemma 5,

$$\boldsymbol{u}^\top \mathrm{Cov}(\boldsymbol{w})\boldsymbol{u} = \mathrm{Var}(\boldsymbol{w}^\top \boldsymbol{u}) \leq \frac{3}{2}\nu^2 \boldsymbol{u}^\top \boldsymbol{P}\boldsymbol{u}.$$

□

Lemma 7. *For all $i \in [n]$*

1. *ϵ_i is sub-Gaussian with $\sigma = 1$, and*
2. *ϵ_i is $(4\sqrt{Aq}, 2/B)$ sub-exponential for any $A, B > 0$ satisfying $e^x < 1 + x + Ax^2$ for $x < B$.*

Proof. For the first claim, note that a random variable bounded in $[a, b]$ is sub-Gaussian with $\sigma^2 = \frac{(b-a)^2}{4}$. To prove second claim, we have

$$\begin{aligned}\exp(\lambda\epsilon_i) &= \tilde{q}\exp(2\lambda(1-\tilde{q})) + (1-\tilde{q})\exp(-2\lambda\tilde{q}) \\ &= \exp(-2\lambda\tilde{q})(1+\tilde{q}(\exp(2\lambda)-1)) \\ &< \exp\left(-2\lambda\tilde{q}\right)\exp\left(\tilde{q}(\exp(2\lambda)-1)\right)) \\ &\leq \exp\left(\tilde{q}A(2\lambda)^2\right),\end{aligned}$$

for $2\lambda < B$. The last step follows from $e^x \leq 1 + x + Ax^2$ for $x < B$. Recall that $\tilde{q} < 2q$, we have

$$\exp(\lambda\epsilon_i) \leq \exp\left(\frac{16qA\lambda^2}{2}\right)$$

for $\lambda < B/2$ as desired. □

Let $\boldsymbol{P}_i$, $i \in [n]$ be orthogonal projection matrices onto the span of $\{\boldsymbol{x}_1, .., \boldsymbol{x}_i\}$, that is,

$$\boldsymbol{P}_i := \boldsymbol{P}_{i-1} + \frac{(\boldsymbol{I} - \boldsymbol{P}_{i-1})\boldsymbol{x}_i\boldsymbol{x}_i^\top(\boldsymbol{I} - \boldsymbol{P}_{i-1})}{\|(\boldsymbol{I} - \boldsymbol{P}_{i-1})\boldsymbol{x}_i\|^2},$$

with $\boldsymbol{P}_0 = 0$.

Lemma 8. *Suppose $A, B > 0$ satisfy $e^x < 1 + x + Ax^2$ for all $x < B$. Let $\sigma^2 := c/2q$, $\nu^2 := 8Ac$ and $\alpha := 2/B$. Then*

1. *If $\boldsymbol{w}_{i-1}$ is $(\sigma, \boldsymbol{P}_{i-1})$ sub-Gaussian, $\boldsymbol{w}_i$ is $(\sigma, \boldsymbol{P}_i)$ sub-Gaussian.*
2. *If $\boldsymbol{w}_{i-1}$ is $(\nu, \alpha, \boldsymbol{P}_{i-1})$ sub-exponential, $\boldsymbol{w}_i$ is $(\nu, \alpha, \boldsymbol{P}_i)$ sub-exponential.*

Proof. We have

$$\begin{aligned}\mathrm{E}\left[\exp(\lambda\boldsymbol{w}_i^\top\boldsymbol{u})\right] &= \mathrm{E}\left[\mathrm{E}\left[\exp(\lambda\boldsymbol{w}_i^\top\boldsymbol{u})\middle|\boldsymbol{w}_{i-1}\right]\right] \\ &= \mathrm{E}\left[\mathrm{E}\left[e^{\lambda\left\langle\boldsymbol{w}_{i-1}, \boldsymbol{u} - 2q\frac{\boldsymbol{x}_i^\top\boldsymbol{u}}{c}\boldsymbol{x}_i\right\rangle + \lambda\epsilon_i\boldsymbol{x}_i^\top\boldsymbol{u}}\middle|\boldsymbol{w}_{i-1}\right]\right] \\ &= \mathrm{E}\left[e^{\lambda\langle\boldsymbol{w}_{i-1}, \boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}\rangle}\mathrm{E}\left[e^{\lambda\epsilon_i\boldsymbol{x}_i^\top\boldsymbol{u}}\middle|\boldsymbol{w}_{i-1}\right]\right].\end{aligned}$$

First, suppose that $\boldsymbol{w}_{i-1}$ is $(\sigma, \boldsymbol{P}_{i-1})$ sub-Gaussian, we will prove that $\boldsymbol{w}_i$ is $(\sigma, \boldsymbol{P}_i)$ sub-Gaussian. By Lemma 7, ϵ_i is 1-sub-Gaussian. Therefore,

$$\begin{aligned}\mathrm{E}\left[e^{\lambda\langle\boldsymbol{w}_{i-1}, \boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}\rangle}\mathrm{E}\left[e^{\lambda\epsilon_i\boldsymbol{x}_i^\top\boldsymbol{u}}|\boldsymbol{w}_{i-1}\right]\right] &\leq \mathrm{E}\left[e^{\lambda\langle\boldsymbol{w}_{i-1}, \boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}\rangle} \cdot e^{\frac{1}{2}\lambda^2\|\boldsymbol{x}_i^\top\boldsymbol{u}\|^2}\right] \\ &\leq e^{\frac{\lambda^2\sigma^2}{2}(\boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u})^\top\boldsymbol{P}_{i-1}(\boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}) + \frac{\lambda^2}{2}(\boldsymbol{u}^\top\boldsymbol{x}_i\boldsymbol{x}_i^\top\boldsymbol{u})}.\end{aligned}$$

We now consider the exponent (divided by the common factor λ^2). It is sufficient to show

$$\begin{aligned}&\frac{\sigma^2}{2}\left(\boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}\right)^\top\boldsymbol{P}_{i-1}\left(\boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}\right) + \frac{1}{2}u^\top\boldsymbol{x}_i\boldsymbol{x}_i^\top\boldsymbol{u} \leq \frac{\sigma^2}{2}\boldsymbol{u}^\top\boldsymbol{P}_i\boldsymbol{u}\ \forall\boldsymbol{u} \\ &\iff \sigma^2\boldsymbol{Q}_{\boldsymbol{x}_i,c}^\top\boldsymbol{P}_{i-1}\boldsymbol{Q}_{\boldsymbol{x}_i,c} + \boldsymbol{x}_i\boldsymbol{x}_i^\top \preceq \sigma^2\boldsymbol{P}_i.\end{aligned}$$

Using Proposition 6 (with $L \leftarrow 1, C \leftarrow c/2q$) and assuming $\sigma^2 = c/2q \geq 1$, we have

$$\begin{aligned}\boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{P}_{i-1}\boldsymbol{Q}_{\boldsymbol{x}_i,c}\sigma^2 + \boldsymbol{P}_{\boldsymbol{x}_i} &= \frac{c}{2q}\left(\boldsymbol{I} - \frac{2q}{c}\boldsymbol{x}_i\boldsymbol{x}_i^\top\right)\boldsymbol{P}_{i-1}\left(\boldsymbol{I} - \frac{2q}{c}\boldsymbol{x}_i\boldsymbol{x}_i^\top\right) + \boldsymbol{x}_i\boldsymbol{x}_i^\top \\ &\preceq \frac{c}{2q}\boldsymbol{P}_i.\end{aligned}$$

Therefore, $\boldsymbol{w}_i$ is a $(c/2q, \boldsymbol{P}_i)$ sub-Gaussian random vector.

Now suppose that $\boldsymbol{w}_{i-1}$ is $(\nu, \alpha, \boldsymbol{P}_{i-1})$ sub-exponential, we will prove that $\boldsymbol{w}_i$ is $(\nu, \alpha, \boldsymbol{P}_i)$ sub-exponential. Again, by Lemma 7, ϵ_i is $(4\sqrt{Aq}, 2/B)$ sub-exponential. Therefore,

$$\mathrm{E}\left[e^{\lambda\langle \boldsymbol{w}_{i-1}, \boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}\rangle}\mathrm{E}\left[e^{\lambda\epsilon_i \boldsymbol{x}_i^\top \boldsymbol{u}}|\boldsymbol{w}_{i-1}\right]\right] \leq \mathrm{E}\left[e^{\lambda\langle \boldsymbol{w}_{i-1}, \boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}\rangle} \cdot e^{8Aq\lambda^2\|\boldsymbol{x}_i^\top \boldsymbol{u}\|^2}\right]$$

for $|\lambda| < \frac{2}{B|\boldsymbol{x}_i^\top \boldsymbol{u}|} = \frac{1}{\alpha\sqrt{\boldsymbol{u}^\top \boldsymbol{x}_i \boldsymbol{x}_i^\top \boldsymbol{u}}}$. Since $\boldsymbol{w}_{i-1}$ is $(\nu, \alpha, \boldsymbol{P}_{i-1})$ sub-exponential,

$$\mathrm{E}\left[e^{\lambda\langle \boldsymbol{w}_{i-1}, \boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}\rangle}\right] \leq e^{\frac{\lambda^2\nu^2}{2}\left(\boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}\right)^\top \boldsymbol{P}_{i-1}\left(\boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}\right)}$$

for $|\lambda| < \frac{1}{\alpha\sqrt{\boldsymbol{u}^\top \boldsymbol{P}_{i-1}\boldsymbol{u}}}$. Note that $\boldsymbol{u}^\top \boldsymbol{P}_i \boldsymbol{u}$ is greater than both $\boldsymbol{u}^\top \boldsymbol{P}_{i-1}\boldsymbol{u}$ and $\boldsymbol{u}^\top \boldsymbol{x}_i \boldsymbol{x}_i^\top \boldsymbol{u}$. We have

$$\mathrm{E}\left[e^{\lambda\langle \boldsymbol{w}_{i-1}, \boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}\rangle}\mathrm{E}\left[e^{\lambda\epsilon_i \boldsymbol{x}_i^\top \boldsymbol{u}}|\boldsymbol{w}_{i-1}\right]\right] \leq e^{\frac{\lambda^2\nu^2}{2}\left(\boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}\right)^\top \boldsymbol{P}_{i-1}\left(\boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}\right) + 8Aq\lambda^2(\boldsymbol{u}^\top \boldsymbol{x}_i \boldsymbol{x}_i^\top \boldsymbol{u})}$$

for all $|\lambda| < \frac{1}{\alpha\sqrt{\boldsymbol{u}^\top \boldsymbol{P}_i \boldsymbol{u}}}$. Hence, it is sufficient to show

$$\begin{aligned}
&\frac{\nu^2}{2}\left(\boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}\right)^\top \boldsymbol{P}_{i-1}\left(\boldsymbol{Q}_{\boldsymbol{x}_i,c}\boldsymbol{u}\right) + 8Aq\boldsymbol{u}^\top \boldsymbol{x}_i \boldsymbol{x}_i^\top \boldsymbol{u} \leq \frac{\nu^2}{2}\boldsymbol{u}^\top \boldsymbol{P}_i \boldsymbol{u}\ \forall \boldsymbol{u} \\
&\iff \nu^2 \boldsymbol{Q}_{\boldsymbol{x}_i,c}^\top \boldsymbol{P}_{i-1}\boldsymbol{Q}_{\boldsymbol{x}_i,c} + 16Aq\boldsymbol{x}_i\boldsymbol{x}_i^\top \preceq \nu^2 \boldsymbol{P}_i.
\end{aligned}$$

This follows from Proposition 6 by substituting $L \leftarrow 16Aq$ and $C \leftarrow c/2q$ and noting that $\nu^2 = 8Ac = Lc$.

□

Lemma 9. *If $c = \min(1/q, 9.3)\log(2n/\delta)$ then with probability at least $1-\delta$, we have*

$$|\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i| < c \text{ for all } i \in [n].$$

Proof. By definition, c is either equal to $\log(2n/\delta)/q$ or $9.3\log(2n/\delta)$. We consider these two cases.

First suppose $c = \log(2n/\delta)/q$. With $\sigma^2 = c/2q$, we have

$$c = \log(2n/\delta)/q = \sigma\sqrt{2\log(2n/\delta)}.$$

By Lemma 3,

$$\begin{aligned}
P\left(|\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i| > c\right) &\leq P\left(|\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i| > c\sqrt{\boldsymbol{x}_i^\top \boldsymbol{P}_{i-1}\boldsymbol{x}_i}\right) \\
&\leq 2\exp\left(-\frac{c^2}{2\sigma^2}\right) \leq \delta/n.
\end{aligned}$$

The result then follows by a union bound over $i \in [n]$.

Now suppose $c = 9.3\log(2n/\delta)$. Note that $A = 0.5803$ and $B = 0.4310$ satisfy $e^x < 1 + x + Ax^2$ for $x < B$. Let $\nu^2 = 8Ac$ and $\alpha = 2/B$. We have

$$\frac{\nu^2}{\alpha} = \frac{8Ac}{2/B} = 4ABc > c.$$

Therefore, by Lemma 4,

$$\begin{aligned}
P\left(|\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i| > c\right) &\leq P\left(|\boldsymbol{w}_{i-1}^\top \boldsymbol{x}_i| > c\sqrt{\boldsymbol{x}_i^\top \boldsymbol{P}_{i-1}\boldsymbol{x}_i}\right) \\
&\leq 2\exp\left(-\frac{c^2}{2\nu^2}\right) = 2\exp\left(-\frac{c}{16A}\right) \\
&< 2\exp\left(-\frac{c}{9.3}\right) \leq \delta/n.
\end{aligned}$$

Again, the result follows by union bounding over $i \in [n]$. □

Lemma 9 and Lemma 8 together prove Theorem 1.

A.2. Proof of Theorem 2

Proof of Theorem 2. Let $D(\delta)$ be the discrepancy obtained by `BinaryBalance` as a function of the failure probability δ. We will show that $(2\log k)D(\delta/k)$ is the corresponding discrepancy obtained by Algorithm 2. Theorem 2 will then follow from Proposition 2. Notice that with probability δ/k, each run of `BinaryBalance` at an internal node of T has discrepancy $D(\delta/k)$. By union bounding over $O(k)$ internal nodes, we have that with probability $1-\delta$, all of the discrepancies are bounded by $D(\delta/k)$.

Assume all the discrepancies at the internal nodes in T are bounded, we show how to bound the discrepancy between any two treatments. Let l and l' be any two leaves in T. The goal is to show that

$$\left\| \frac{\alpha(l')}{\alpha(l')+\alpha(l)} s(l) - \frac{\alpha(l)}{\alpha(l')+\alpha(l)} s(l') \right\|_\infty$$

is small. First we relate $s(v)$ to $s(v_l)$ and $s(v_r)$ where v_l, v_r are the left and right children of v. By definition,

$$w(v) = \frac{\alpha(v_r)}{\alpha(v_l)+\alpha(v_r)} s(v_l) - \frac{\alpha(v_l)}{\alpha(v_l)+\alpha(v_r)} s(v_r)$$

and

$$s(v) = s(v_l) + s(v_r).$$

Therefore,

$$w(v) = s(v_l) - \frac{\alpha(v_l)}{\alpha(v_l)+\alpha(v_r)} s(v),$$

and

$$-w(v) = s(v_r) - \frac{\alpha(v_r)}{\alpha(v_l)+\alpha(v_r)} s(v).$$

Hence, both

$$\left\| s(v_l) - \frac{\alpha(v_l)}{\alpha(v)} s(v) \right\|_\infty \quad \text{and} \quad \left\| s(v_r) - \frac{\alpha(v_r)}{\alpha(v)} s(v) \right\|_\infty$$

are bounded by $D(\delta/k)$.

Now consider v_1, v_2 and v_3 in T such that v_1 is a child of v_2 and v_2 is a child of v_3. We have, by triangle inequality,

$$\left\| s(v_1) - \frac{\alpha(v_1)}{\alpha(v_3)} s(v_3) \right\|_\infty \le \left\| s(v_1) - \frac{\alpha(v_1)}{\alpha(v_2)} s(v_2) \right\|_\infty + \frac{\alpha(v_1)}{\alpha(v_2)} \left\| s(v_2) - \frac{\alpha(v_2)}{\alpha(v_3)} s(v_3) \right\|_\infty \le \left(1 + \frac{\alpha(v_1)}{\alpha(v_2)}\right) D(\delta/k).$$

Let l be a leaf in T and let $lv_1v_2\dots r$ be the path from l to the root r. Repeatedly applying the above relation along the path gives

$$\left\| s(l) - \frac{\alpha(l)}{\alpha(r)} s(r) \right\|_\infty \le \left(1 + \frac{\alpha(l)}{\alpha(v_1)} + \frac{\alpha(l)}{\alpha(v_2)} \dots + \frac{\alpha(l)}{\alpha(r)}\right) D(\delta/k). \tag{3}$$

Since there are at most $\log k$ nodes in the path from l to r,

$$\left\| s(l) - \frac{\alpha(l)}{\alpha(r)} s(r) \right\|_\infty \le (\log k) D(\delta/k).$$

Finally, for any two leaves l and l',

$$\left\| \frac{s(l)/\alpha(l) - s(l')/\alpha(l')}{1/\alpha(l) + 1/\alpha(l')} \right\|_\infty \le \left\| \frac{s(l)/\alpha(l) - s(r)/\alpha(r)}{1/\alpha(l) + 1/\alpha(l')} \right\|_\infty + \left\| \frac{s(l')/\alpha(l') - s(r)/\alpha(r)}{1/\alpha(l) + 1/\alpha(l')} \right\|_\infty \le (2\log k) D(\delta/k).$$

□

Remark 2. *If all weights are uniform, the summation in (3) becomes a geometric series and can be bounded by a constant. Therefore, we can remove the factor* $\log k$ *in Theorem 2.*

A.3. Other Proofs

Proof of Proposition 1. Let $\boldsymbol{B} = \begin{bmatrix} \sqrt{\phi}\boldsymbol{I} \\ \sqrt{1-\phi}\boldsymbol{X}^\top \end{bmatrix}$. We have from Threorem 1 that $\boldsymbol{w}_n = \boldsymbol{B}\boldsymbol{\eta} = \begin{bmatrix} \sqrt{\phi}\boldsymbol{\eta} \\ \sqrt{1-\phi}\boldsymbol{X}^\top\boldsymbol{\eta} \end{bmatrix}$ is a mean zero $(c/2q, \boldsymbol{P})$ sub-Gaussian random vector, where

$$\begin{aligned}\boldsymbol{P} &= \boldsymbol{B}(\boldsymbol{B}^\top\boldsymbol{B})^{-1}\boldsymbol{B}^\top \\ &= \begin{bmatrix} \phi\boldsymbol{Q} & * \\ * & * \end{bmatrix}.\end{aligned}$$

Therefore, by sub-Gaussianity of $\boldsymbol{w}_n$, for any vector $\boldsymbol{u}$, we have

$$\mathrm{E}\left[\exp\left(\sqrt{\phi}\boldsymbol{\eta}^\top\boldsymbol{u}\right)\right] \leq \exp\left(\frac{c}{4q}\boldsymbol{u}^\top(\phi\boldsymbol{Q})\,\boldsymbol{u}.\right)$$

We therefore have the sub-Gaussian claim. The sub-exponential result follows similarly. □

Proof of Proposition 2. Suppose $c = \log(2n/\delta)/q$. We have from Proposition 1 that $\boldsymbol{z}$ is a $(\sqrt{c/2q}, \boldsymbol{Q})$ sub-Gaussian vector. This implies that $\mathrm{Cov}(\boldsymbol{z}) \preceq \frac{c}{2q}\boldsymbol{Q}$.

When $c = 9.3\log(2n/\delta)$, we have from Proposition 1 that $\boldsymbol{\eta}$ is a $(\sqrt{8Ac}, \alpha, \boldsymbol{Q})$ sub-exponential vector. Now, Lemma 5 gives that

$$\mathrm{Cov}(\boldsymbol{z}) \preceq \frac{3}{2}8Ac\boldsymbol{Q} = 12Ac\boldsymbol{Q}.$$

□

Proof of Proposition 2. When $c = \frac{\log(4n/\delta)}{q}$ we have that with probability at least $1-\delta/2$, $\boldsymbol{B}\boldsymbol{\eta}$ is a $\left(\sqrt{c/2q}, \boldsymbol{Q}\right)$ sub-Gaussian vector. Since $\boldsymbol{B}\boldsymbol{\eta} = \begin{bmatrix} \sqrt{\phi}\boldsymbol{\eta} \\ \sqrt{1-\phi}\boldsymbol{X}^\top\boldsymbol{\eta} \end{bmatrix}$ and $\boldsymbol{Q} \preceq \phi\boldsymbol{I}$, we have $\sqrt{1-\phi}\sum_i \eta_i\boldsymbol{x}_i = \sqrt{1-\phi}\boldsymbol{X}^\top\boldsymbol{\eta}$ is a $\left(\sqrt{1/\phi}\sqrt{c/2q}, \boldsymbol{I}\right)$ sub-Gaussian random vector. By sub-Gaussian concentration, we have with probability at least $1-\delta/2d$, $|\left(\boldsymbol{X}^\top\boldsymbol{\eta}\right)^\top\boldsymbol{e}_i| \leq \left(\sqrt{(c/2\phi(1-\phi)q}\right)\sqrt{4d/\delta}$. The result follows by a union bound over $\boldsymbol{e}_1, ..., \boldsymbol{e}_d$ and $\|\boldsymbol{X}^\top\boldsymbol{\eta}\|_2 \leq \sqrt{d}\|\boldsymbol{X}^\top\boldsymbol{\eta}\|_\infty$.

When $c = 9.3\log(4n/\delta)$, then with probability $1-\delta$, $\begin{bmatrix} \sqrt{\phi}\boldsymbol{\eta} \\ \sqrt{1-\phi}\boldsymbol{X}^\top\boldsymbol{\eta} \end{bmatrix}$ is a $(\sqrt{8Ac}, \alpha, \boldsymbol{Q})$ random vector. Like before, this implies $\sqrt{1-\phi}\boldsymbol{X}^\top\boldsymbol{\eta}$ is a $(\sqrt{8Ac}, \alpha, \phi\boldsymbol{I})$ random vector. By sub-exponential concentration, we have

$$P\left(\left|\sqrt{1-\phi}\left(\boldsymbol{X}^\top\boldsymbol{\eta}\right)^\top\boldsymbol{e}_i\right| \geq t\right) \leq \begin{cases} 2\exp\left(-t^2\phi/2\nu^2\right) & \text{when } t \leq \frac{\nu^2}{\alpha\sqrt{\phi}} \\ 2\exp\left(-t/2\alpha\right) & \text{otherwise} \end{cases}$$

Setting $t = \sqrt{2\log(4d\delta)(8Ac)} \leq c \leq \nu^2/\alpha$, (when $n \geq d$), we get that with probability at least $1-\delta/2d$, we have

$$|\sqrt{1-\phi}\left(\boldsymbol{X}^\top\boldsymbol{\eta}\right)^\top\boldsymbol{e}_i| \leq 9.3\sqrt{\frac{\log(4d/\delta)\log(4n/\delta)}{\phi}}.$$

The rest follows by a union bound. □

Proof of Proposition 3. First consider the case when $c = \log(2n/\delta)/q = \sigma\sqrt{2\log(2n/\delta)}$. By Proposition 1, $\boldsymbol{\eta}$ is $(\sigma, \boldsymbol{Q})$ sub-Gaussian with $\sigma = \sqrt{c/2q}$. From Lemma 3, we have

$$P(|\boldsymbol{\eta}^\top\boldsymbol{\mu}| > c\sqrt{\boldsymbol{\mu}^T\boldsymbol{Q}\boldsymbol{\mu}}) \leq 2\exp\left(-\frac{c^2}{2\sigma^2}\right) = \delta/n.$$

Now consider $c = 9.3\log(2n/\delta)$. By Proposition 1, $\boldsymbol{\eta}$ is $(\nu, \alpha, \boldsymbol{Q})$ sub-exponential with $\nu = \sqrt{8Ac}$. Note that

$$\nu^2/\alpha = 8Ac/(2/B) = 4ABc > c.$$

From Lemma 4, we have

$$\begin{aligned} P(|\boldsymbol{\eta}^\top \boldsymbol{\mu}| > c\sqrt{\boldsymbol{\mu}^T \boldsymbol{Q}\boldsymbol{\mu}}) &\leq 2\exp\left(-\frac{c^2}{2\nu^2}\right) \\ &= 2\exp\left(-\frac{c}{16A}\right) \leq \delta/n. \end{aligned}$$

The result then follows by a union bound. □

Proof of Proposition 5. This follows from Proposition 2 and the proof of Theorem 3 in (Harshaw et al., 2020). □

A.4. Robustness

Proposition 2 immediately gives a bound on $\lambda_{\max}(\mathrm{Cov}(\boldsymbol{z}))$ and hence bounds accidental bias.

Remark 3 (Accidental Bias)**.** *With probability* $\geq 1 - \delta$ *the maximum eigenvalue of* $\mathrm{Cov}(\boldsymbol{z})$ *satisfies*

$$\lambda_{\max}\left(\mathrm{Cov}(\boldsymbol{z})\right) \leq \frac{\sigma^2}{\phi}.$$

B. Simulations

B.1. Description

DGP Name	$\mathbf{X}$	$\mathbf{y(0)}$	$\mathbf{y(a)}$ s.t. $\mathbf{a \neq 0}$
QuickBlockDGP	$X_{i,k} \sim \mathcal{U}(0,10), \forall k \in \{1,2\}$	$\prod_{k=1}^{2} X_k + \epsilon$	1 + y(0)
LinearDGP	$X_{i,k} = \epsilon_k, k \in \{1,\ldots,4\}$	$\mathbf{X}\beta + \frac{1}{10}\epsilon_{y(0)}$	1 + $\mathbf{X}\beta + \frac{1}{10}\epsilon_{y(1)}$
LinearDriftDGP	$X_{i,k} = \frac{i}{N} + \epsilon_k, k \in \{1,\ldots,4\}$	$\mathbf{X}\beta + \frac{1}{10}\epsilon_{y(0)}$	1 + $\mathbf{X}\beta + \frac{1}{10}\epsilon_{y(1)}$
LinearSeasonDGP	$X_{i,k} = \sin(2\pi\frac{i}{N}) + \epsilon_k, k \in \{1,\ldots,4\}$	$\mathbf{X}\beta + \frac{1}{10}\epsilon_{y(0)}$	1 + $\mathbf{X}\beta + \frac{1}{10}\epsilon_{y(1)}$
QuadraticDGP	$X_{i,k} = 2\beta_k - 1, k \in \{1,2\}$ $\mu_0 = X_1 - X_2 + X_1^2 + X_2^2 - 2X_1X_2$	$\mu_0 + \frac{1}{10}\epsilon_{y(0)}$	1 + $\mu_0 + \frac{1}{10}\epsilon_{y(1)}$
CubicDGP	$X_{i,k} = 2\beta_k - 1, k \in \{1,2\}$ $\mu_0 = X_1 - X_2 + X_1^2 + X_2^2 - 2X_1X_2$ $+X_1^3 - X_2^3 - 3X_1^2X_2 + 3X_1X_2^2$	$\mu_0 + \frac{1}{10}\epsilon_{y(0)}$	1 + $\mu_0 + \frac{1}{10}\epsilon_{y(1)}$
SinusoidalDGP	$X_{i,k} = 2\beta_k - 1, k \in \{1,2\}$ $\mu_0 = \sin\left(\frac{\pi}{3} + \frac{\pi X_1}{3} - \frac{2\pi X_2}{3}\right)$ $-6\sin(\frac{\pi X_1}{3} + \frac{\pi X_2}{4}) + 6\sin(\frac{\pi X_1}{3} + \frac{\pi X_2}{6})$	$\mu_0 + \frac{1}{10}\epsilon_{y(0)}$	1 + $\mu_0 + \frac{1}{10}\epsilon_{y(1)}$

Table A1. Data generating processes used in simulations. All ϵs indicate a standard normal variate and all βs indicate a standard uniform variate. i indicates a unit's index. Covariate vectors are row-normalized to unit norm, except for the QuickBlock simulation which just normalized relative to the maximum row norm.

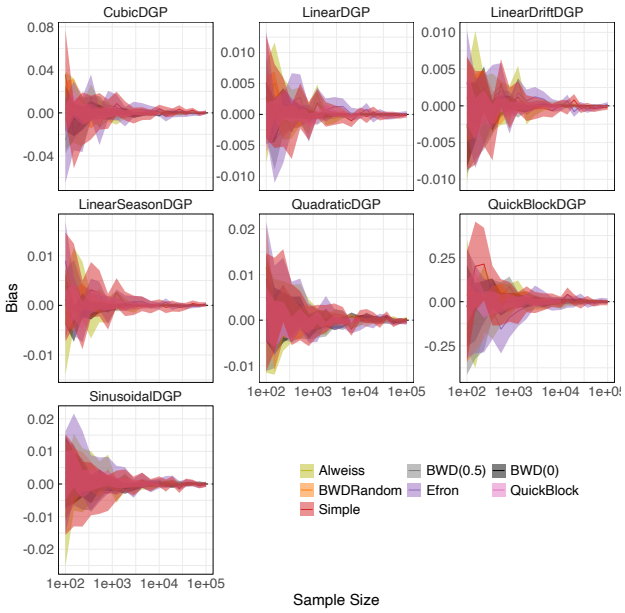

Figure A1. Bias

B.2. Binary Treatments

Bias. Figure A1 shows that none of the examined methods are biased (but that does not imply that they are robust (Efron, 1971; Harshaw et al., 2020)).

Imbalance. To measure linear imbalance, we calculate the L_2 norm of the difference in covariate means. While this is far from the only measure of imbalance, it serves as an effective metric to demonstrate how well various metrics serve to eliminate linear imbalances, a common diagnostic used my experimenters. We further normalize across sample-size by multiplying by the sample size. Since all methods are unbiased, this has the effect of showing parametric convergence rates as a flat line in the graph. Unsurprisingly, methods which directly optimize for linear imbalance perform very well in Figure A2. Our proposed algorithm has better finite sample imbalance minimization than does the algorithm of (Alweiss et al., 2020) due to the finite sample improvements of (Dwivedi and Mackey, 2021).

B.3. Non-uniform assignment

MSE. Figure A3 shows the resulting mean squared-error attained by methods which support marginal probabilities not equal to one-half. All methods perform well, with DM performing effectively on nearly linear processes, and QuickBlock performing slightly better when the true process is highly non-linear.

Marginal probability. The evaluation in Figure A4 examines how closely each method hews to the desired marginal probability of treatment. All methods do a good job of ensuring the appropriate marginal distribution.

B.4. Multiple Treatments

Entropy. To ensure that treatment is being assigned with the correct marginal probability, we can measure the normalized entropy of the empirical marginal treatment probabilities. If the values are perfectly uniform, then the value will be exactly one. As the normalized entropy decreases, the marginal probabilities are more uneven, indicating a failure to match the desired marginal distribution. BWD performs very similarly to complete randomization (which almost exactly matches the desired marginal probabilities), slightly out-performing (Smith, 1984). Note that while Smith (1984) only seeks to optimize the marginal probability of treatment for each unit, BWD additionally balances covariates.

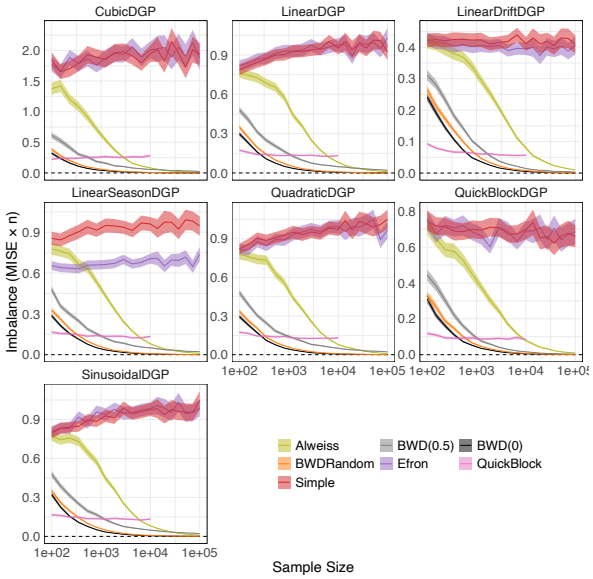

Figure A2. Imbalance. BWD is highly effective at eliminating linear imbalance between groups.

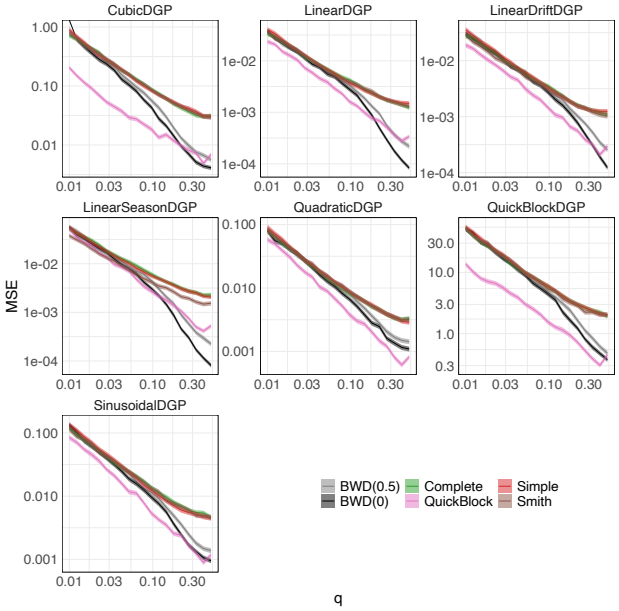

Figure A3. Marginal probability of treatment

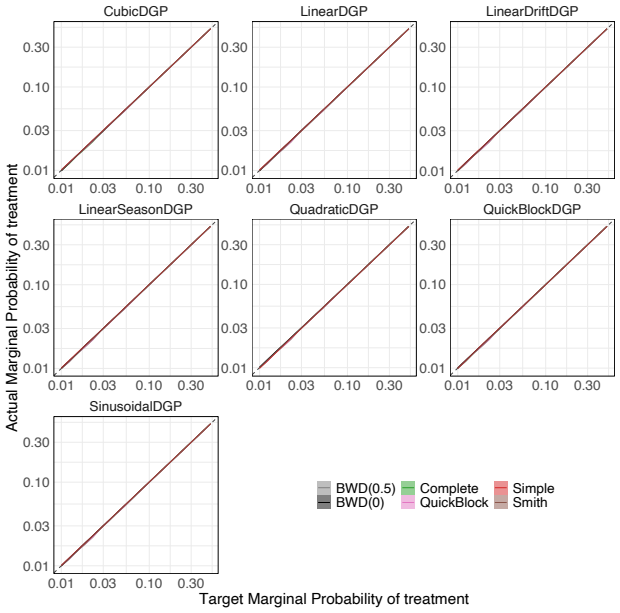

Figure A4. Marginal probability of treatment

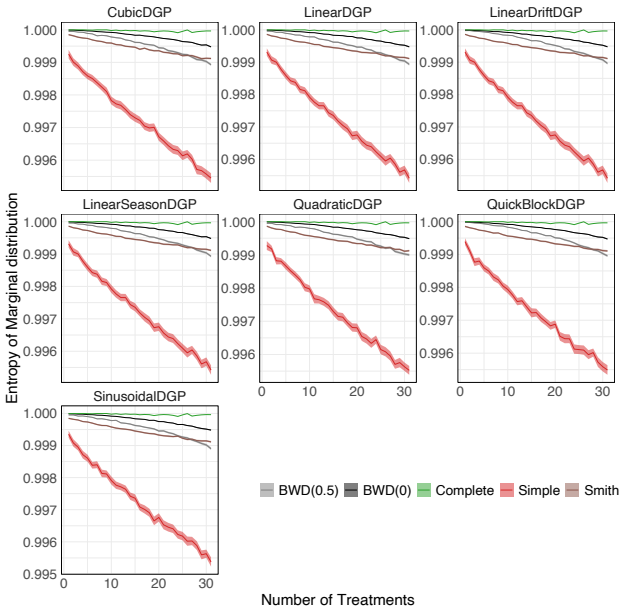

Figure A5. Entropy