Experimental Design with Scientific Applications

Jennifer Brennan

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Kevin Jamieson, Chair

Lalit Jain

Aleksandr Aravkin

Program Authorized to Offer Degree:

Paul G. Allen School of Computer Science and Engineering

University of Washington

Abstract

Experimental Design with Scientific Applications

Jennifer Brennan

Chair of the Supervisory Committee:
Professor Kevin Jamieson
Paul G. Allen School of Computer Science & Engineering

Experimentation is a powerful tool to understand and optimize the world around us. The scientific method, with its emphasis on experimentation, has become the de facto means to generate knowledge across the physical, biological, and increasingly the social sciences. Experimentation is used to determine the functions of genes, design chemical substances with desired properties, evaluate the performance of new medicines, optimize products on the internet, and forecast the potential impacts of economic and social policies.

Every experiment includes an *experimental design*, which specifies the data to be collected. A thoughtful experimental design collects the data that is most informative to the scientific question at hand, allocating the data collection budget to answer the scientific question accurately and efficiently. The importance of careful experimental design is self-evident in long, expensive experiments such as human clinical trials, in which every human subject must be justified from a cost and an ethics perspective. Even in experiments with lower marginal cost, such as high-throughput biological screening or internet A/B testing, choosing the appropriate experimental design can be the difference between making a scientific discovery or losing that signal in the experimental noise.

This thesis addresses several aspects of experimental design. In Chapter 2 we ask how precise our measurements should be when that precision comes at a cost, with applications to the design and analysis of pilot experiments in the setting of high-throughput screening. In Chapters 3 and 4 we ask which of many units to measure if the units all have observed features, with applications to the optimization of antibiotic combinations and estimation in global health. In Chapter 5 we ask how to experiment on a collection of units when experimenting on one unit affects the outcomes of the other units, with applications to A/B testing in online marketplaces.

Table of Contents

Acknowledgments

This thesis would not have been possible without the support of the many people who have shaped my personal, intellectual, and professional development.

First and foremost I would like to thank my advisor, Kevin, for the time and effort he dedicated to mentoring me. Kevin's enthusiasm for solving difficult and impactful research problems is contagious, and some of my favorite moments in grad school were the nights we stayed late at the office to work on a particularly tricky proof. Kevin taught me how to approach a research problem, how to see the big picture without getting caught in the weeds, and how to critically read the literature. Kevin cares both about doing good research and developing good researchers, and I am fortunate to have been in his group.

I have also been fortunate to have a cast of outstanding mentors in graduate school. Lalit Jain pushed me to think with a new degree of mathematical precision, which was painful at times but which has had an enormous impact on my ability to engage with the statistical machine learning literature. Lalit is enthusiastic about sharing his knowledge and is equally excited to discuss a new result, a theorem he has known for years, or a cute math fact. I have also benefited from collaborations at UW and beyond, which have giving me insight into the broader process of scientific research. Sasha Aravkin introduced me to the world of global health modeling and the statistical challenges of quantifying population health. Zach Tatlock gave me an appreciation for the value of simple solutions to complex problems. Erik Wright encouraged me to engage deeply with the biology literature and gave me the tools to design and test my own scientific hypotheses. Jean Pouget-Abadie helped me navigate the balance between theoretical advances and the practical constraints of an experimentation platform. Having the opportunity to work with so many different groups, and across a variety of disciplines, has given me a broader perspective on research as a whole. I have also benefited from other mentors at UW, including Sandy Kalpan, whose editorial review improved several of my papers, and both Jamie Morgenstern and Marina Meila, who asked thoughtful questions as members of my thesis committee.

One of the best parts of grad school has been talking to people who are spending years of their lives immersing themselves in a particular subject. I would like to thank Steven Lyubomirsky, Marisa Kirisame, Altan Haan, and Mike He for inviting me to work on their project Dynamic Tensor Rematerialization, which introduced me to a programming languages perspective on deep learning; Swati Padmanabhan for being my local expert in convex optimization; Robbie Weber for discussions on algorithms and pedagogy; John

Thickstun for discussions on everything ML-related, including a course he taught on generative modeling; Kendall Lowrey, for keeping me on my toes with his skepticism of probability; Josue Nassar and Ben Evans, who along with Kendall invited a frequentist to join their Bayesian project; Aravind Rajeswaran, who co-mentored me for my first research project; Peter Ney, who shared many stories about the security challenges of curating genetic data; Gabe Erion, who shared his interest in feature attribution methods; and Ann Donnelly, who carefully, patiently, and enthusiastically explained the microbiology of antibiotics to a computer scientist.

I would also like to thank the people who helped me get to grad school: Prof Jessica Wu of Harvey Mudd College, who advised my undergraduate research; Prof Ran Libeskind-Hadas, who went out of his way to support my graduate school application; and Prof Michael Orrison, who helped me clarify my thoughts around the choice of attending graduate school or going to industry. Andrew Fishberg also provided valuable advice when I was considering graduate school, and continues to help me figure out what really matters in life. I would also like to thank Aaron Archer, who has provided me with thoughtful advice, most notably the encouragement to get a research internship in industry during graduate school.

Finally, my family has been a constant source of support. I would like to thank my parents for empowering me to set my own goals, helping me develop the tools to pursue them, and celebrating when I accomplished them. I would also like to thank my husband Eamon, who helps me remember the big picture.

Chapter 1

Introduction

An *experimental design* describes the data points that will be collected and analyzed to evaluate a scientific hypothesis. As an example, consider the variety of experiments that might be performed to discover and evaluate a new antibiotic. A first step in drug discovery may be to screen candidate molecules; for example an entire library of potential drugs may each be tested against *E. coli* to identify molecules that inhibit bacterial growth. After a likely candidate has been identified, scientists may test it at a variety of doses, in different growing conditions, or in combination with other drugs to understand how the drug may behave in the human body. Years later, when the antibiotic is deployed in a clinical setting, public health officials may evaluate the antibiotic not only as a cure for a single patient but also as part of the larger healthcare system. Experiments could be used to understand the system-level effects of prescribing this drug, such as the emergence of resistant mutants.

A well-designed experiment allows the experimenter to answer a scientific question with the greatest accuracy and in the fewest measurements possible. Although the experiments described above differ in scale, purpose, and methods, they all share common design principles. At the heart of a good experimental design is an understanding of the *observation model*, which describes the (usually probabilistic) mechanism by which nature generates data each time a measurement is taken. Knowing the observation model allows the scientist to choose measurements that are most informative for the question at hand. In this thesis we make contributions to experimental design for observation models at the three levels of complexity described in our example above: a simple model in which many similar experiments are run in parallel (the small molecule screen); a more complex model in which known variables may affect the outcome (the tests in different growing conditions); and finally a systems-level model in which different units of experimentation may interact with one another, so that outcomes are no longer independent (understanding the emergence of resistance).

We elaborate on each level of complexity and our contributions in the following sections.

1.1 A simple model: identical parallel experiments

Perhaps the simplest of all scientific experiments is one in which a value is measured and compared to a known standard. These experiments are ubiquitous in science and society. Process engineers measure the frequency of defective widgets in a batch to ensure that the company's average production quality meets a published standard. Schools administer standardized tests to determine whether students are reading at grade level. Marine biologists test the oysters on a particular

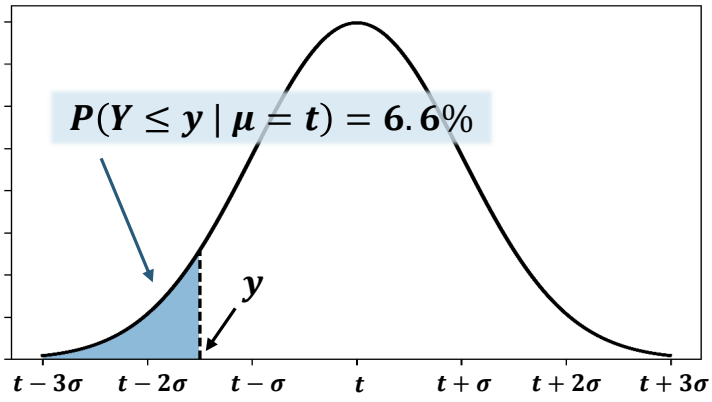


Figure 1.1: A test for the null hypothesis $H_0: \ \mu = t$ versus the alternate hypothesis $H_1: \ \mu < t$ when observation (bacterial growth) is distributed $Y \sim \mathcal{N}(\mu, \sigma)$. Suppose the measured growth for a certain candidate drug was $y = t - 1.5\sigma$. If the candidate actually had no effect on bacterial growth, then there is a 6.6% chance that we would observe measured growth at least this low purely by chance (the area of the blue shaded region is 6.6% of the total area under the curve). As a result, we would fail to reject the null hypothesis at the 5% level - that is, we cannot claim that this candidate inhibits growth to level t or lower.

beach to determine their levels of saxitoxin, which causes paralytic shellfish poisoning in humans when safe levels are exceeded.

In our drug discovery example, an experiment that screens a library of small molecules for their antibiotic activity against *E. coli* might first measure the growth of the bacteria when a specific concentration of each small molecule is added. The growth under a given drug candidate could then be compared to a predefined threshold that indicated whether the candidate was effective enough to merit further study. The observed growth is a random variable that we want to compare to the threshold. The statistical field of *hypothesis testing* gives us tools to make this comparison while taking into account the randomness in the measured outcome. Once we know the observation model that generated the growth measurements (i.e., the noise model on the measurements) we can determine the probability that, on average, applying this molecule to *E. coli* would result in growth below the threshold value. Figure 1.1 illustrates the principles of hypothesis testing when the observation model is additive Gaussian noise. For a detailed introduction to hypothesis testing, see Chapter 10 of Wasserman's *All of Statistics* (2004).

If we took multiple independent observations (replicates) of growth for each drug candidate, we would reduce the noise σ in our observation model, thereby increasing our chance of identifying promising drugs. Choosing the number of replicates to perform is an important part of experimental design. This choice involves trading off between the accuracy gain and the expense of additional observations. In Chapter 2 we introduce a method for choosing the number of replicates based on

the result of an inexpensive (low-replicate) pilot experiment. Our method works by providing a high-probability uniform lower bound on the cumulative distribution function of a mixing distribution after observing a draw of the mixture, making advances in statistics to improve the design of experiments.

1.2 A more complex model: adding features

Additional complexity arises when the observations depend on known variables, or *features*, of the unit that is being measured. For example, in agricultural science the crop yield could be modeled as a function of the fertilizer, light, and water a plant receives. The efficacy of a new drug for hypertension may depend on a patient's age or comorbidities. The strength of concrete depends on the relative ratios of its component parts, as well as the ambient temperature and humidity as it cures.

An observation model for these scenarios typically takes the form $Y_i = f(X_i) + \varepsilon_i$, where $f(X_i)$ gives the expected outcome $\mathbb{E}[Y_i]$ and ε_i is a mean-zero additive noise term often modeled using the normal distribution $\varepsilon_i \sim \mathcal{N}(0, \sigma)$. For example, if $f(X_i) = \theta^T X_i$ for some fixed but unknown *parameter vector* θ, then the observation model is the statistical linear model.

A central question, and one that is often synonymous with the term "experimental design," is the following: which units X_i should you choose to measure, given a limited budget, to most accurately estimate f? Approaches to optimal experimental design often rely on the *Fisher information matrix*, which captures the amount of information a set of measurements $\{X_i\}_{i=1}^n$ contains about the parameters θ of f. If f is a linear model then the Fisher information is solely a function of the measurements $\{X_i\}_{i=1}^n$, not the observations $\{Y_i\}_{i=1}^n$, so that the optimal design can be computed before any observations are made. If the model f is nonlinear in X then the optimal choice of measurements may depend on the true parameter θ, in which case multiple rounds of design and estimation are required. The topic of experimental design in the linear setting is explored in more detail in Pukelsheim's classic textbook *Optimal Design of Experiments* (2006), while strategies for the nonlinear case are discussed in Pronzato and Pázman's *Design of Experiments in Nonlinear Models* (2013).

In this thesis we consider two observation models that incorporate features. In Chapter 3, we seek the best combination of antibiotics to use against *E. coli*. We model the effect of an antibiotic combination as a function that maps a vector of concentrations (specifying a given dosing of each drug) to a binary outcome (presence or absence of growth in *E. coli*). The experimental decision in this case is the choice of which of the exponentially many concentration vectors to test. We identify biologically meaningful constraints on the observation model f that allow us to identify the lowest-dose combination that suppresses bacterial growth without having to search the entire space of concentration vectors. We use our design to measure the interactions between all combinations of eight antibiotics, finding no interactions of clinically relevant magnitudes among the experimental settings tested.

In Chapter 4, we consider models for under-reported count data motivated by the incomplete reporting of birth defects worldwide. We use the Fisher information matrix to demonstrate the

statistical challenges of estimating counts from under-reported data, and illustrate that without further constraints the problem is statistically difficult. We demonstrate the use of constraints and regularization to improve the count estimates, providing applications to the under-reporting of medical care when only hospital inpatient data is recorded.

1.3 Experimentation in an interconnected system: interference between observations

When an experiment takes measurements from multiple units, the unit-level observations can be thought of as one vector $\mathbf{Y}$. Each student in a class may be randomly assigned to levels of an educational intervention, and their measured reading scores form the vector $\mathbf{Y}$. A new environmental remediation may be performed on several streams, with $\mathbf{Y}$ representing the change in heavy metal concentration of each stream. The participants of a clinical trial may be assigned to either receive a drug or a placebo, and $\mathbf{Y}$ records their health outcomes.

A common assumption is that the distribution of $\mathbf{Y}$ can be factored as the product of the distributions of the N outcomes Y_i, that is, that the observations are statistically independent. When this assumption of independence is violated, the traditional tools of experimental design are more difficult to apply.

Consider a public health intervention to prescribe a new antibiotic in hospital settings as a means to reduce mortality due to infections. The choice of antibiotic is made for each patient, and researchers are interested in understanding the difference in outcome (fatal infections) between two universes: (1) when all patients are treated with the new antibiotic, and (2) when all patients are treated with the old antibiotic. In the language of randomized controlled trials, the new antibiotic would be the "treatment" condition while the old antibiotic would be the "control." The estimand of interest, which determines whether a large-scale migration to the new antibiotic provides enough benefit to justify such a change, is the *average treatment effect* τ:

$$\tau = \frac{1}{N} \sum_{i \in [N]} \mathbb{E}[Y_i | \mathbf{Z} = \mathbf{1}] - \mathbb{E}[Y_i | \mathbf{Z} = -\mathbf{1}].$$

In this notation, the condition $\mathbf{Z} = \mathbf{1}$ indicates that all N units are assigned the treated condition; this is universe (1) described above. Since each patient can only be chosen to receive one antibiotic (new or old), we cannot simply observe the outcomes under the two universes (complete treatment or complete control) and compare them. Instead, we must choose to assign some patients to the treatment, some to the control, and use the resulting per-patient outcomes to estimate the effect of the intervention.

A naive randomized controlled trial in this setting might independently assign each patient to treatment ($Z_i = 1$) or control ($Z_i = -1$) with some probability, observe the outcomes Y_i, and then report the difference-in-means estimate

$$\widehat{\tau} = \frac{1}{|\{i : Z_i = 1\}|} \sum_{i:Z_i=1} Y_i - \frac{1}{|\{i : Z_i = -1\}|} \sum_{i:Z_i=-1} Y_i.$$

Unfortunately, in many real-world examples this approach will fail due to *interference* between units. In the antibiotics example, treating one patient with a more effective drug is likely to provide protection for other patients in their hospital by reducing the background level of bacteria circulating in the hospital. Alternatively, if the bacteria responded to the new drug by evolving to become more virulent, the prescription of the drug to one patient could have negative impacts on the patients in the control condition who catch the mutated bug. In either case, neither the treatment nor the control units of our randomized trial are representative of their behavior under the universal treatment or control condition. To accurately estimate the average treatment effect τ in the presence of such interference we have to change the estimator, change the experimental design, or both.

In Chapter 5 of this thesis we consider experimental design in *bipartite networks* subject to interference. In the bipartite setting, units on one side of the network are assigned to treatment or control, but their outcomes are influenced by the assignments of other units through each unit's interaction with the other side of the bipartite network. Examples include marketplaces where buyers interact with items and digital platforms where service providers interact with users.

An important experimental design in the network setting is the *cluster-randomized design*, in which units are first clustered and then a treatment status is assigned to each cluster. We propose a novel clustering objective for the bipartite setting that produces a balanced partitioning of the experimental units based on the degree to which each unit's assignment influences the outcome of other units. We show that this objective minimizes the bias of the difference in means estimator in the setting of interference under a linear model on the outcomes Y_i. Furthermore, we show that the bias reduction is robust to the choice of model for Y_i, and support our theoretical results with empirical studies.

1.4 Organization of this thesis

This thesis is based on work conducted with a variety of research collaborators, with whom the author has been fortunate to work during the past five years. Each chapter describes a distinct project.

- Chapter 2 introduces a new method for power analysis using pilot experiments, and is based on the paper "Estimating the number and effect sizes of non-null hypotheses," (Brennan, Vinayak, and Jamieson, 2020), which appeared at the International Conference on Machine Learning.

- Chapter 3 proposes, analyzes, and implements a new experimental design for provably identifying synergistic antibiotic combinations using only a fraction of the measurements required by traditional methods. It is based on the paper "Sample-Efficient Identification of High-Dimensional Antibiotic Synergy with the Normalized Diagonal Sampling Design" (Brennan, Jain, Garman, Donnelly, Wright, and Jamieson, 2022), published in PLOS Computational Biology.

- Chapter 4 discusses the challenges of modeling under-reported count data, suggests methods for improving the fit of these models using constraints and regularization, and presents applications to global health data. This work was completed as part of a collaboration at the Institute for Health Metrics and Evaluation in Seattle, and a preprint is available on Arxiv with the title "Analysis and Methods to Mitigate the Effect of Under-reporting in Count Data" (Brennan, Bannick, Kassebaum, Wilner, Thomson, Aravkin, and Zheng, 2021).

- Chapter 5 analyzes cluster-randomized designs in the setting of network interference for bipartite experiments. It is based on work completed while the author was an intern at Google Research, and is currently under review. Jean Pouget-Abadie and Vahab Mirrokni from Google also contributed to this work.

The author also had the opportunity to participate in other research projects during her PhD in which she was not the primary author; these works are not discussed in this thesis, but are briefly described here:

- Working with members of the UW Programming Languages and Software Engineering group (PLSE), we developed and analyzed a method to train larger deep learning models with a limited memory budget by evicting intermediate activations to trade space for time – see "Dynamic Tensor Rematerialization" (Kirisame, Lyubomirsky, Haan, Brennan, He, Roesch, Chen, and Tatlock, 2021), published at the 2021 International Conference on Learning Representations.

- In collaboration with other students in UW ML, we provided a precise definition of audio-to-score alignment in the domain of automatic music transcription (see "Rethinking evaluation methodology for audio-to-score alignment," (Thickstun, Brennan, and Verma, 2020)).

- Working with collaborators from UW and Stony Brook, we enabled Bayesian methods to adapt to a changing environment via selective memory of past observations (see "BAM: Bayes with Adaptive Memory," (Nassar, Brennan, Evans, and Lowrey, 2022), published at the 2022 International Conference on Learning Representations).

Chapter 2

Estimating the Number and Effect Sizes of Non-null Hypotheses

In this chapter, we study the problem of estimating the distribution of effect sizes in a multiple testing setting. Knowing this distribution allows us to calculate the power of any experimental design. We show that it is possible to estimate this distribution using an inexpensive pilot experiment, which takes significantly fewer samples than would be required by an experiment that identified the discoveries. Our estimator can be used to guarantee the number of discoveries that will be made using a given experimental design in a future experiment. We prove that this simple and computationally efficient estimator enjoys a number of favorable theoretical properties, and demonstrate its effectiveness on data from a gene knockout experiment on influenza inhibition in *Drosophila*.

2.1 Introduction

Designing scientific experiments is something of a chicken and egg problem. In order to design an experiment with a specified power (type II error), we need to know the effect size (the mean of the test statistic under the alternate hypothesis). The effect size determines the required accuracy of each measurement, which increases with the number of *experimental replicates* (samples). Unfortunately, this effect size is typically unknown, and estimating the effect size for a single hypothesis test is as sample intensive as performing the original experiment. In the case of single hypothesis testing, this presents a fundamental barrier to efficient experimental design. By contrast, in the setting of multiple hypothesis testing, we show that it is possible to estimate the distribution of effect sizes present in the data using an inexpensive pilot experiment, which takes significantly fewer samples than would be required for the full experiment.

For example, suppose a scientist would like to test 10,000 genes using an experimental measurement that is distributed $\mathcal{N}(\mu_i, \frac{1}{t})$ when the effect size is μ_i and t replicates are performed. Without knowledge of the likely effect sizes, it is unclear how to choose an experimental design. An experiment with too many replicates per hypothesis is wasteful; one with too few will lack the statistical power to identify alternate hypotheses. In this chapter, we seek to facilitate experimental design in the multiple testing setting by answering the question *"How many hypotheses have an effect size of at least γ?"* using significantly fewer samples than would be needed to identify all discoveries with that effect size. These estimates suggest a trade-off between the cost of an experiment (as measured by the number of experimental replicates required to achieve a certain

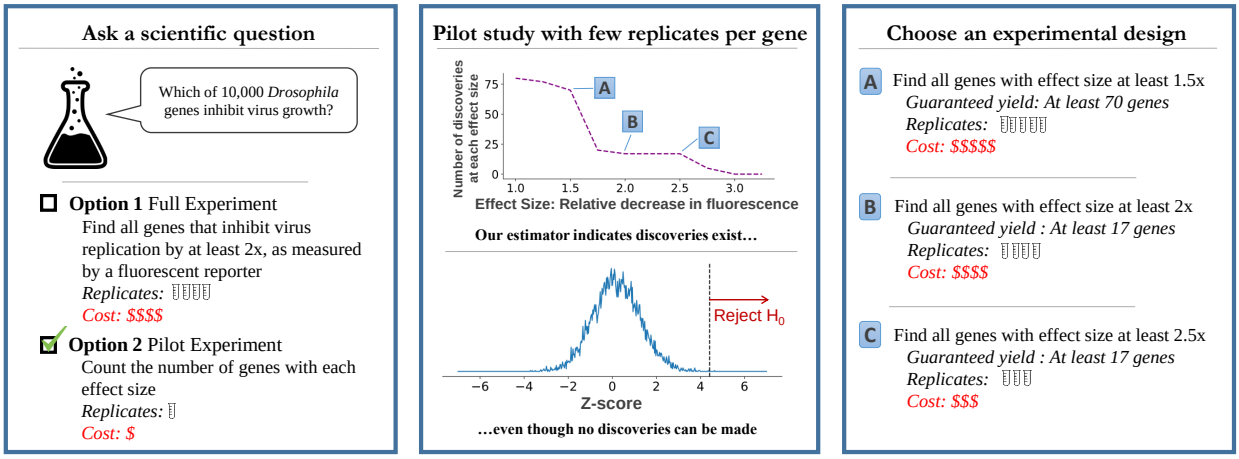


Figure 2.1: When applied to the results of a pilot experiment, our estimator can estimate the cost and number of discoveries guaranteed by different experimental designs. In this example, the original experiment design (Option 1) is expensive, with no guarantee on the number of discoveries that will be made. Our method suggests two alternatives to the original experimental design (B); the same guarantee on discoveries could be made at lower cost (C), or additional discoveries could be made at higher cost (A).

power) and the number and effect sizes of the discoveries that will be made. Figure 2.1 illustrates the application of our estimator to an inexpensive pilot study, allowing a scientist to evaluate possible experimental designs. The application to experimental design motivates an important property of our estimator: it must produce a *conservative* estimate of the number of hypotheses above a given effect size. If the scientist designs a costly experiment based on the results of this estimator, it is important to ensure that this experiment will generate at a minimum the estimated number of discoveries.

As a baseline, one approach to this estimation problem is to use a *plug-in estimator*, which estimates the entire distribution of effect sizes and then "plugs in" this estimate as if it were the true distribution. The plug-in estimator could start with the maximum likelihood estimate (MLE) of the distribution of effect sizes given the observed test statistics. The estimate for the fraction of hypotheses above some effect size γ would simply be the fraction of this distribution that exceeded γ. Unfortunately, such a plug-in estimator based on the MLE may vastly overestimate this fraction, as two distributions can have similar likelihoods but very different amounts of mass above some threshold, as illustrated in Figure 2.2.

In this chapter, we design an estimator for the fraction of hypotheses with effect sizes above a given threshold, for all thresholds simultaneously. Our estimator operates in the spirit of the Kolmogorov-Smirnov test, first creating an ℓ_∞ ball around the empirical CDF to define plausible distributions, and then finding the element of the ball with the smallest amount of probability

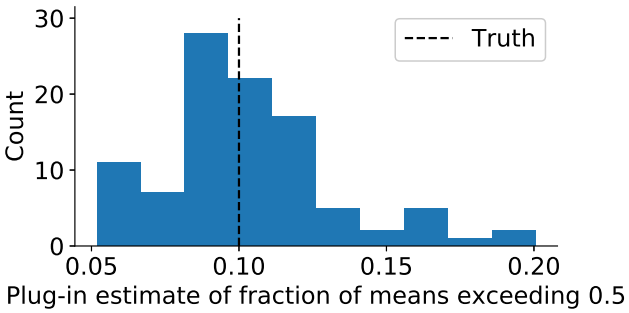


Figure 2.2: This histogram shows the results of 100 simulations in which $n \approx 10^5$ samples were drawn from a Gaussian mixture model $\frac{9}{10}\mathcal{N}(0,1) + \frac{1}{10}\mathcal{N}(1,1)$. The MLE plug-in estimator was used to estimate the fraction of observations with mean greater than $\frac{1}{2}$. The MLE is frequently over-optimistic about the number of discoveries to be made, misleading an experimenter into expecting more discoveries than are possible.

mass above γ. With high probability, this amount of mass does not exceed the true fraction of hypotheses with mean at least γ. We prove that this simple and computationally efficient estimator enjoys a number of favorable theoretical properties, including finite-sample upper and lower bounds on the value of the estimate.

2.1.1 Problem Statement

Let ν_* be a distribution on $\mathbb{R}$, and for $i = 1, 2, \ldots n$ let

$$\mu_i \sim \nu_*$$

be an unobserved latent variable drawn iid from ν_*. For each μ_i drawn from ν_*, we observe the test statistic

$$X_i \sim f_{\mu_i},$$

where f_μ is a known distribution parameterized by the effect size μ. For example, if the test statistics are Z-scores, we would have $f_\mu = \mathcal{N}(\mu, 1)$. While our estimator is well defined for any parametric f (e.g., any single-parameter exponential family), we focus on Gaussian test statistics for exposition. In the setting of Figure 2.1, ν_* represents the distribution of effect sizes and X_i are the observations.

Our goal is to estimate the probability that the effect size of an observation is greater than γ,

$$\zeta_{\nu_*}(\gamma) := \mathbb{P}_{\nu_*}(\mu > \gamma), \tag{2.1}$$

simultaneously for all $\gamma \in \mathbb{R}$.

The problem of counting the non-null hypotheses is most interesting when γ is small. For example, consider the case when the test statistics X_i are Z-scores. Under the hypothesis that all effect sizes are zero, the expected maximum Z-score is $\mathbb{E}[\max_i X_i] \approx \sqrt{\log n}$. Therefore, if we want to avoid any false discoveries, we cannot reject any hypotheses with test statistic less than $\Theta(\sqrt{\log n})$. If the effect sizes are at least this large, then we will be able to identify the alternate hypotheses through a standard Bonferroni correction (Dunn, 1961). In this regime, counting is no more difficult than identification. However, if the effect sizes are much smaller than this threshold (say, if all $\mu_i \ll 1$), identification could be impossible. Our estimator, by contrast, detects the existence of discoveries even in this low signal-to-noise regime.

2.1.2 Contributions

Our contributions are as follows:

- Given a parameterization f_μ, we propose an estimator that provides a conservative estimate of the fraction of effect sizes above a given threshold, simultaneously for all thresholds (Section 2.2).
- We provide finite-sample bounds on the error of our estimator (Theorem 2.2.1).
- In the low signal-to-noise regime and the setting of Gaussian mixtures, we compare our estimator's sample complexity to known lower bounds for hypothesis testing (detecting the presence of the alternate hypothesis), and we give a novel lower bound for the sample complexity of estimation (estimating the fraction of means from the alternate hypothesis). We show that our method matches finite-sample and asymptotic rates for these problems, even though it is designed for more general distributions than the ones in these lower bounds (Section 2.2).
- We describe how to use this estimator to design pilot studies for scientific experimentation (Section 2.3). When testing n hypotheses in the low signal-to-noise regime, our technique detects treatments with positive effect sizes using a factor of n fewer replicates than it would take to identify them. Additionally, the results of the pilot experiment can be used to upper bound the cost of identifying the discoveries at each effect size.

2.1.3 Related Work

The problem of estimating the number of null hypotheses has been studied extensively in the statistics literature. Our goal in this chapter is to provide a conservative estimate of the number of hypotheses with effect size above some threshold (Eqn (2.1)). There are several lines of work related to this goal.

Simple Null Hypotheses A different but related problem is to estimate the number of non-null hypotheses, regardless of their effect sizes, i.e., $\mathbb{P}_{\nu_*}(\mu \neq 0)$. In this setting - also known as the simple null hypothesis - it is possible to compute p-values that are uniformly distributed under the null. For example, when observations are drawn $X_i \sim \mathcal{N}(\mu_i, 1)$, the p-value is $p_i = 1 - \Phi(X_i)$, where Φ is the standard normal CDF.

The graphical estimator of Schweder & Spjøtvoll (1982) was the first technique to estimate the number of nulls, using the principle that p-values are distributed uniformly under the null hypothesis and skewed toward zero under the alternate. Their technique estimates the density of the p-value distribution at 1. This same idea was improved in the context of estimating the number of nulls for adaptive control of the false discovery rate (FDR) (Benjamini and Hochberg, 2000; Storey, 2002). These later works provide finite-sample guarantees on overestimating the number of nulls in order to make non-asymptotic guarantees on FDR control. However, none of these results provide lower bounds on the estimated number of non-nulls. Motivated by adaptive FDR control, techniques for counting the number of non-null hypotheses have been extended to incorporate prior knowledge about the dependence structure of the hypotheses or the likelihood that each test will result in a discovery. See Li & Barber (2019) for a review of this area.

Another technique for the simple null setting, again motivated by the uniform distribution of p-values under the null, is to test the extent to which the distribution of p-values deviates from the uniform distribution. Several estimators have taken this approach (Genovese et al., 2004; Meinshausen and Rice, 2006; Patra and Sen, 2016; Jin, 2008). Most similar to our work are the techniques that build one-sided confidence intervals around the empirical CDF of p-values (Genovese et al., 2004; Meinshausen and Rice, 2006), which provide finite-sample error bounds and a conservative estimator. Finally, there are estimators specific to the Gaussian setting, which estimate the zero-mean component in a mixture of Gaussians (Cai et al., 2007; Carpentier and Verzelen, 2019).

Extensions to one-sided null hypotheses ($H_0 : \ \mu \leq 0$) further assume that p-values are subuniformly distributed when $\mu < 0$ (Meinshausen and Bühlmann, 2005; Li and Barber, 2019) or assume a gap between 0 and the smallest alternate effect size (Lee and Valiant, 2021). These works estimate the quantity $\mathbb{P}_{\nu_*}(\mu > 0)$. This problem is a special case of ours, because subuniformity holds only for the threshold of $\gamma = 0$.

Composite Null Hypotheses We seek to estimate the number of hypotheses with an effect size above some threshold. Here, p-values are neither subuniform nor necessarily well defined, so much of the previous work is not applicable. The Fourier transform technique (Jin, 2008) can be extended to address composite null hypotheses (Chen, 2019). However, this extension only provides asymptotic results, which are insufficient since we seek a conservative estimator.

Adapting the Generalized Likelihood Ratio Test Jiang & Zhang (2016) develop asymptotic power statements for the generalized likelihood ratio test for Gaussian observations. We discuss in Section 2.5 how this work could be used to create an estimator for our problem, and highlight the limitations that make this approach impractical.

Plug-in Estimation As discussed in Section 2.1, another approach to this problem is plug-in estimation, where an estimate $\widetilde{\nu}$ of the distribution ν_* is used to form an estimator $\widehat{\zeta_n}(\gamma) = \mathbb{P}_{\widetilde{\nu}}(\mu > \gamma)$. When f_μ is Gaussian, the task is to learn a mixture of Gaussians. In this setting, much effort has been devoted to recovering the mixture parameters (Pearson, 1894; Belkin and Sinha, 2010; Kalai et al., 2010; Hardt and Price, 2015) or learning a mixture that is close to the original distribution in some metric, such as total variation (TV) distance (Moitra and Valiant, 2010; Daskalakis and Kamath, 2014). Outside of the Gaussian setting, recent works have provided

guarantees for learning mixtures of binomial distributions in terms of the Wasserstein-1 distance (Tian et al., 2017; Vinayak et al., 2019). These types of theoretical guarantees do not lend themselves easily to guarantees on our problem, since two distributions can be close in TV or Wasserstein distance but have very different amounts of mass above some threshold γ.

Empirical Bayes Methods Our estimator takes advantage of multiple hypothesis testing by using the empirical distribution of the X_i to learn something about the latent distribution ν_*. The same idea can be seen in empirical Bayes methods, where the empirical distribution of X_i is used as the prior over X. Several papers have taken an empirical Bayes approach to multiple testing, but none address our exact problem. Efron (2007) uses an empirical Bayes method to estimate the distribution of X under the alternate hypothesis, which is distinct from our goal of estimating ν_* (note we cannot simply deconvolve Efron's estimate to get ν_*, as it is not guaranteed to have any parametric form). Stephens (2017) uses empirical Bayes methods and a strong unimodality assumption on ν_* to produce estimates and confidence intervals for each μ_i. While these confidence intervals could theoretically be used to estimate (2.1), the fact that Stephens' method produces a confidence interval for individual μ_i suggests that they will be too loose to compete with our method. Furthermore, this method only works for Gaussian and t-distributed observations.

2.2 Estimating effect sizes

Recall our goal, to estimate $\zeta_{\nu_*}(\gamma)$ from Eqn (2.1). Let $\widehat{F}_n(t) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}\{X_i \leq t\}$ be the empirical CDF of the test statistics X_i and

$$F_\nu(t) = \mathbb{P}_{\mu\sim\nu,\ X\sim f_\mu}(X \leq t)$$

be the true CDF of test statistics under latent distribution ν. For any $\gamma \in \mathbb{R}$, our estimator is given by

$$\widehat{\zeta}_n(\gamma) = \min_{\nu:||\widehat{F}_n - F_\nu||_\infty \leq \tau_{\alpha,n}} \int_\gamma^\infty \nu(x)dx \tag{2.2}$$

where the estimator is conservative with probability at least $1-\alpha$, and

$$\tau_{\alpha,n} = \sqrt{\frac{\log(2/\alpha)}{2n}}.$$

The intuition for this estimator is as follows. To conservatively estimate the amount of mass ζ above threshold γ, we look for the distribution with the smallest amount of mass above γ that could have plausibly generated the observations X_i. Our measurement of plausibility is based on high probability bounds on the deviation between the empirical CDF and its expectation. If $\widehat{F}_n$ was in fact drawn from F_ν, then with high probability the ℓ_∞ distance between $\widehat{F}_n$ and F_ν will not exceed $\tau_{\alpha,n}$. By restricting our search space to the ℓ_∞ ball around $\widehat{F}_n$ (seen in the constrained optimization from Eqn (2.2)), we do not overestimate the true amount of mass above γ, with high probability. Moreover, using different values of γ traces a curve for $\zeta_{\nu_*}(\gamma)$ (see the middle panel of

Figure 2.1). We note that this estimator can be implemented as an efficient convex program. We simply discretize x over some range, and the estimator becomes a convex program in the vector $\mathbf{x}$. It can then be solved using off-the-shelf software (see Appendix A.4 for details).

2.2.1 Main Results

Our estimator underestimates the true mass $\zeta_{\nu_*}(\gamma)$ for all γ simultaneously with high probability. Furthermore, we provide a finite sample bound on how much we underestimate $\zeta_{\nu_*}(\gamma)$ at every γ.

Theorem 2.2.1. *For $i = 1, \ldots, n$, let $\mu_i \sim \nu_*$ and $X_i \sim f_{\mu_i}$ where each draw is iid. Let our simultaneous estimator be given by (2.2). Then, the probability of overestimating the fraction of hypotheses with effect size above any threshold γ is bounded by α:*

$$\mathbb{P}\left(\exists \gamma : \widehat{\zeta}_n(\gamma) > \zeta_{\nu_*}(\gamma)\right) \leq \alpha.$$

Furthermore, with probability at least $1-\delta$, for all $\gamma \in \mathbb{R}$ and $\varepsilon \in (0, \zeta_{\nu_*}(\gamma)]$ we have $\zeta_{\nu_*}(\gamma) - \hat{\zeta}_n(\gamma) \leq \varepsilon$ whenever

$$n \geq \frac{\log\left(\frac{4}{\alpha\delta}\right)}{\left(\min_{\nu:\mathbb{P}_\nu((\gamma,\infty)) \leq \zeta_{\nu_*}(\gamma)-\varepsilon} ||F_\nu - F_{\nu_*}||_\infty\right)^2}. \tag{2.3}$$

This guarantee, that our estimator is conservative, is critical in the use of pilot studies to guide experimental design. The key quantity in this sample complexity result is the minimum ℓ_∞ distance between the true CDF F_{ν_*} and the set of CDFs corresponding to mixing distributions with less than ζ mass above γ. We call this set of mixing distributions S,

$$S(\zeta, \gamma) := \{\nu : \mathbb{P}_\nu((\gamma, \infty)) \leq \zeta\}. \tag{2.4}$$

Specifically, consider $S(\zeta_{\nu_*}(\gamma) - \varepsilon, \gamma)$, which appears in Eqn (2.3). If $\varepsilon = 0$, then we have $\nu_* \in S(\zeta_{\nu_*}, \gamma)$, so the minimum ℓ_∞ distance to F_{ν_*}, $\min_{\nu \in S(\zeta_{\nu_*}(\gamma),\gamma)} ||F_\nu - F_{\nu_*}||_\infty$, would be zero, implying that no finite sample can guarantee $\varepsilon = 0$. This reflects the fact that $\widehat{\zeta}_n(\gamma)$ is an underestimate at every γ; therefore, in order for ε to be zero, we must have estimated $\zeta_{\nu_*}(\gamma)$ exactly. As ε increases, $S(\zeta_{\nu_*}(\gamma) - \varepsilon, \gamma)$ shrinks, and the distance to F_{ν_*} increases, decreasing the required number of samples n.

To interpret the sample complexity in Theorem 2.2.1, we consider a simple model where test statistics are drawn from a mixture of two Gaussians. In this setting, which we denote $X_i \sim P(\zeta_*, \gamma_*)$, we have

$$\begin{aligned} \mu_i &\sim (1-\zeta_*)\delta_0 + \zeta_* \delta_{\gamma_*} \\ X_i &\sim \mathcal{N}(\mu_i, \sigma^2), \end{aligned} \tag{2.5}$$

where δ_x is the Dirac delta function at x. There are two natural questions we might ask: How many samples are necessary to determine the existence of the mixture component at $\gamma_* > 0$, and how many samples are required to estimate the weight of this component? We call these the *testing* and *estimation* problems respectively. In the following sections, we address our algorithm's sample complexity for the testing and estimation problems, and compare to lower bounds. For ease of exposition, let $\alpha = \delta$, although the results hold for the more general case.

2.2.2 The Testing Problem

In the testing problem, we observe X_i according to (2.5), and we want to determine whether $\zeta_* > 0$ (i.e., testing $H_0 : \ P_{\nu_*}(\mu > 0) = 0$ vs $H_1 : \ P_{\nu_*}(\mu > 0) > 0$). Our test declares H_1 if $\widehat{\zeta}_n(0) > 0$, and H_0 if $\widehat{\zeta}_n(0) = 0$. Clearly this test erroneously declares H_1 with probability at most δ (it has type I error at most δ), since $\widehat{\zeta}_n(0) \leq \zeta_*$ with probability at least $1 - \delta$. The next corollary bounds the sample complexity that guarantees a probability of detection of at least δ (i.e., that bounds the type II error by δ).

Corollary 2.2.2. *Let $\{X_i\}_{i=1}^n$ be drawn according to* (2.5)*. Consider the simultaneous estimator $\widehat{\zeta}_n$ defined by (2.2). Then, with probability at least $1 - \delta$, we have $\widehat{\zeta}_n(0) \leq \zeta_*$ and $\widehat{\zeta}_n(0) > 0$ whenever*

$$n \geq \frac{2 \log\left(\frac{2}{\delta}\right)}{\zeta_*^2 \left(\Phi_\sigma\left(\frac{1}{2}\gamma_*\right) - \Phi_\sigma\left(-\frac{1}{2}\gamma_*\right)\right)^2},$$

where Φ_σ is the CDF of the distribution $\mathcal{N}(0, \sigma^2)$. Furthermore, if $\gamma_* < \sigma$, then the above can be simplified to

$$n \geq \frac{16\sigma^2 \log\left(\frac{2}{\delta}\right)}{\zeta_*^2 \gamma_*^2}.$$

Comparison to finite sample lower bounds

The finite-sample lower bound arises from the "most biased coin problem" (Chandrasekaran and Karp, 2014; Jamieson et al., 2016). In this problem, the algorithm draws N observations X_i as per (2.5), where N is potentially a random variable, according to either $H_0 : X_i \sim \mathcal{N}(0, \sigma^2)$ or $H_1 : X_i \sim P(\zeta_*, \gamma_*)$. When γ_* and ζ_* are known and $\gamma_* \leq \sigma$, Theorem 2 of Jamieson et al. (2016) states that any (potentially randomized) procedure that decides between these hypotheses with probability of error at most δ requires at least

$$\mathbb{E}[N] \geq \max\left\{\frac{1-\delta}{\zeta_*}, \ \frac{\sigma^2 \log(1/\delta)}{2\zeta_*^2 \gamma_*^2}\right\}$$

samples. To facilitate comparison with the sample complexity of our estimator, we show in Lemma A.2.3 that the small-γ_* sample complexity from Corollary 2.2.2 matches the stated lower bound up to constants both when δ is fixed and as $\delta \to 0$.

Comparison to asymptotic lower bounds

Next, we compare our estimator to known results for the hypothesis test deciding between $H_0 : X_i \sim \mathcal{N}(0, 1)$ and $H_1 : \ X_i \ \sim P(\zeta_n, \gamma_n)$, where the alternate hypothesis is parameterized by

$$\zeta_n = n^{-\beta}, \ \ \gamma_n = \sqrt{2r \log n} \tag{2.6}$$

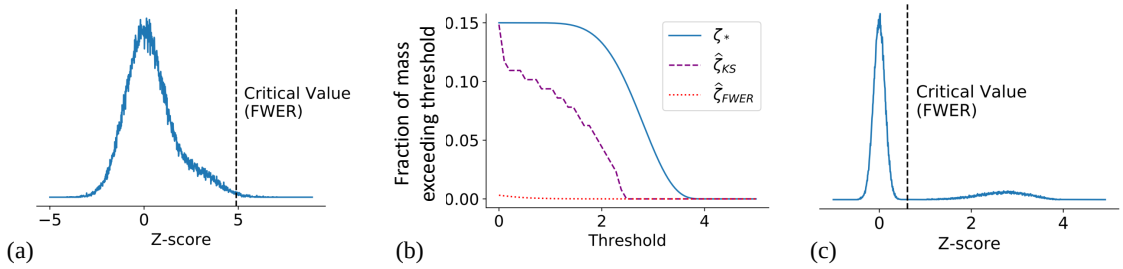


Figure 2.3: Our estimator applied to pilot studies. (a) After observing $X_i \sim \mathcal{N}(\mu_i, 1)$ for $i = 1, \ldots, n$ with $n = 10^4$, only 0.3% of null hypotheses are rejected via a Bonferroni corrected test (indicated by the FWER critical value). However, the Z-scores appear skewed positive, suggesting additional discoveries exist. (b) Our estimator $\widehat{\zeta}_{KS}$ indicates that there are many discoveries to be made; for example, at least 9% of treatments have effect size at least 1, and at least 4% have effect size at least 2. Note that our estimator also counts more discoveries at each threshold than are identified by Bonferroni correction ($\widehat{\zeta}_{FWER}$), without exceeding the true value ζ_*. (c) The experimenter designs an experiment to identify the effects greater than 2, and allocates $\gamma^{-2} \log(n) \log(1/\widehat{\zeta}(\gamma)) = 8$ replicates per hypothesis. Now, 14% of the null hypotheses can be rejected.

and $\sigma^2 = 1$. The higher criticism work of Donoho & Jin (2004) defines various regimes of β. In the so-called *dense regime*, when $\beta \in (0, \frac{1}{2})$, the alternate is detectable (i.e., there is a test with asymptotically vanishing type 1 and type 2 errors) for all $r > 0$. Our test has the same power in this regime, as stated in the following lemma.

Proposition 2.2.3. *Consider the hypothesis test defined by* $H_0 : X_i \sim \mathcal{N}(0, 1)$ *and* $H_1 : X_i \sim P(\zeta_n, \gamma_n)$*, parameterized by* (2.6) *with* $\sigma = 1$*. If* $0 < \beta < \frac{1}{2}$*, then our test has asymptotically full power in the detectable region, i.e.,*

$$\lim_{n \to \infty} \mathbb{P}_0(\{\textit{reject } H_0\}) + \mathbb{P}_1(\{\textit{fail to reject}\}) = 0$$

for any $r > 0$*.*

We emphasize that our test matches lower bounds in this highly restricted setting of known one- or two-spike Gaussians, even though our estimator is more general and does not assume the given hypothesis parameterization.

We also note that the dense regime (many small effects) is of interest for our setting; we are interested in detecting arbitrarily small effect sizes, which is possible in the dense regime but not in the sparse regime. For detection in the sparse regime, that is, when $\beta > 1/2$, the effect size must be large, on the order of $\Theta(\sqrt{\log n})$ (Donoho and Jin, 2004). Section 2.5 describes a potential modification to this test using an empirical Bernstein bound, which may enable it to detect in the sparse regime.

2.2.3 The Estimation Problem

In the estimation problem, we observe X_i according to (2.5), and we estimate ζ_* using our estimator $\widehat{\zeta}(0)$. Since $\widehat{\zeta}(0) \leq \zeta_*$ with high probability, it remains to understand the magnitude of this underapproximation $-$ the dependence of ε from Theorem 2.2.1 on the number of samples n. The following corollary describes the number of samples needed to guarantee an error bound $\varepsilon \leq \frac{1}{2}\zeta_*$ with high probability.

Corollary 2.2.4. *Let $X_i \sim \mathcal{N}(\mu_i, 1)$ and $\mu_i \sim \nu_*$, with $\nu_* = (1-\zeta_*)\delta_0 + \zeta_*\delta_{\gamma_*}$. Let $\zeta_* > 0$ and $\gamma_* \in (0, \sigma]$. Then, with probability at least $1-\delta$, our estimate $\widehat{\zeta}_n$ from (2.2) satisfies $\widehat{\zeta}_n(0) \in (\frac{1}{2}\zeta_*, \zeta_*]$ as long as*

$$n \gtrsim \frac{\sigma^4 \log\left(\frac{2}{\delta}\right)}{\zeta_*^2 \gamma_*^4}.$$

We present a novel lower bound for the estimation problem which matches our result up to constants, the proof of which can be found in Appendix A.3.2.

Lemma 2.2.5. *Consider data $\{X_i\}_{i=1}^n$ generated under the model $X_i \sim P(\zeta, \gamma)$, parameterized by $\zeta \in (0, \frac{1}{2})$ and $\gamma \in (0, \sigma)$ according to our canonical two-spike model (2.5). Fix a parameterization (ζ_*, γ_*). For any $\varepsilon \in (0, \frac{2}{3}\zeta_*)$, define the set A_ε of nearby parameterizations as*

$$A_\varepsilon = \left\{(\zeta, \gamma) \ : \ |\zeta_* - \zeta| \leq 4\varepsilon, \ \tfrac{1}{3}\gamma_* \leq \gamma \leq 3\gamma_*\right\}.$$

Suppose $\widehat{\zeta}_n(X)$ is an estimator of ζ satisfying $\mathbb{P}\left(|\widehat{\zeta}_n(X) - \zeta| \geq \varepsilon\right) < \frac{1}{4}$ for any $(\zeta, \gamma) \in A_\varepsilon$. Then the estimator requires at least $n \gtrsim \frac{\sigma^4}{\varepsilon^2 \gamma_*^4}$ samples on the instance (ζ_*, γ_*).

Instantiating Lemma 2.2.5 with $\varepsilon = \frac{1}{2}\zeta_*$, we see that the sample complexity in Corollary 2.2.4 matches the lower bound.

Finally, we remark that estimating the fraction of observations with mean γ_* to constant multiplicative error requires a factor of $\frac{\sigma^2}{\gamma_*^2}$ more samples than testing whether any observations have mean γ_*.

2.3 Applications to pilot experiments

In this section, we consider the task of analyzing a pilot experiment to count the number of discoveries when the effect sizes are small, say $\mu < 1$ for the case of Gaussian $\mathcal{N}(\mu, 1)$ observations. Figure 2.3 illustrates this process. First, the scientist allocates a small number of replicates to a large number of hypotheses in order to obtain many noisy estimates of effect sizes (Fig. 2.3(a)). The scientist then uses our estimator to obtain a guarantee on the number of discoveries to be made at each effect size (Fig. 2.3(b)). Finally, the scientist calculates the cost of identifying the discoveries that have been detected using a choice of fixed and sequential experimental designs.

When the full experiment is run, it results in at least as many discoveries as our estimator has guaranteed (Fig. 2.3(c)).

The following proposition describes our estimator's performance on pilot data in the low signal-to-noise regime. In particular, if the pilot study design allocates its replicates equally across all n hypotheses, our estimator detects the alternate hypotheses using a factor of n fewer total replicates than it would take to identify these discoveries.

Proposition 2.3.1. *Consider a pilot experiment for n hypotheses, where an initial budget of $B = mt$ will be used to uniformly allocate t replicates to each of $m \leq n$ randomly chosen hypotheses. Suppose the true distribution of effect sizes is $\nu_* = (1 - \zeta_*)\delta_0 + \zeta_*\delta_{\gamma_*}$ and $f_\mu = \mathcal{N}(\mu, \frac{1}{t})$, as when computing Z-scores from t replicates. Assume that $\gamma_* < \frac{12}{\sqrt{t}}$. Then,*

$$\mathbb{P}(\widehat{\zeta_n}(0) > 0) \geq 1 - \delta$$

i.e., we detect the presence of positive effects with high probability, as long as

$$\gamma_* \geq 4\sqrt{\frac{\log(\frac{2}{\delta})}{\zeta_*^2 B}} \qquad \textit{and} \qquad m \geq \frac{4\log(\frac{2}{\delta})}{\zeta_*^2}.$$

Remark 2.3.1. *Consider the setting where the budget is constrained, say $B \lesssim n$, and let ζ_* be constant (so that the proportion of discoveries does not vanish with n). Proposition 2.3.1 suggests maximizing m: either taking $m = n$ if $B \geq n$ or taking $t = 1$ if $B < n$. With this budget allocation, our estimator detects the existence of alternate effects with just $B \approx \log(\frac{1}{\delta})\gamma_*^{-2}$ total replicates. Note that distinguishing observations from $\mathcal{N}(0, 1)$ and $\mathcal{N}(\gamma_*, 1)$ with probability $1-\delta$ requires $\log(\frac{1}{\delta})\gamma_*^{-2}$ samples, so identifying all of the discoveries requires at least $n\log(\frac{1}{\delta})\gamma_*^{-2}$ replicates total. We conclude that, in this instance, any identification procedure requires n times more total replicates than our estimator requires for detection.*

Since our estimator provides information about the number of hypotheses exceeding each effect size γ, it can also be used to estimate the cost of identifying discoveries with different experimental designs. The experimenter can choose between either a fixed experimental design (in which case, the number of replicates in the design determines both the cost and the discoveries that will be made), or a sequential design as in Jamieson & Jain (2018). A sequential design can be more difficult to implement, but could result in significant savings if the effect sizes are very diverse (for example, if some effects are large and others are small). Our estimator can quantify the advantage of using a sequential experimental design.

2.4 Experiments

Details of our implementation can be found in Appendix A.4.

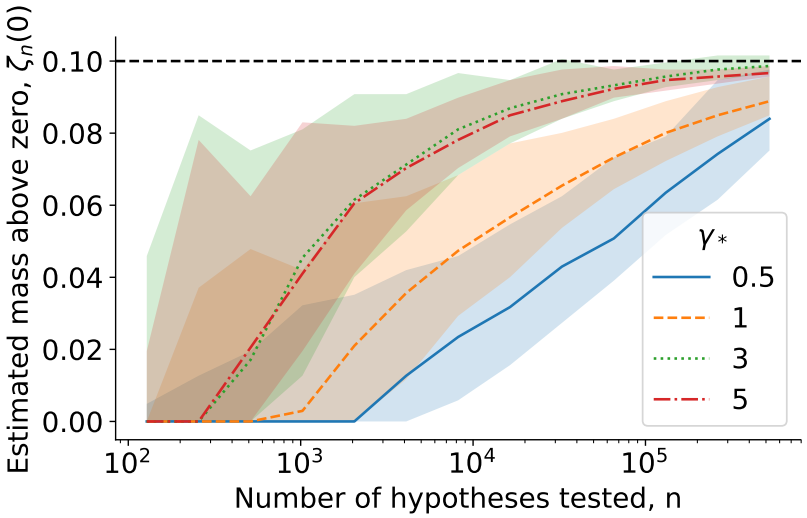


Figure 2.4: Median and 90% bootstrapped confidence intervals for $\widehat{\zeta}_n(0)$, where $\nu_* = (1-\zeta_*)\delta_0 + \zeta_*\delta_{\gamma_*}$, for various γ_*. As n increases, for a constant $\zeta_* = 0.1$, our estimator $\widehat{\zeta}_n(0)$ converges to ζ_* without overestimating. As expected, the estimates are lower (have more error) when the alternate effect size γ_* is small.

2.4.1 Experimental Results on Simulated Data

We evaluate our estimator on both real and simulated data. We begin with the mixture of two Gaussians $\nu_* = (1-\zeta_*)\delta_0 + \zeta_*\delta_{\gamma_*}$. Figure 2.4 shows the rate of convergence of our estimator for different values of the effect size under the alternate, γ_*. Note that the estimate never exceeds the true value ζ_*, and that it improves as n increases. The variance of our estimator, shown with bootstrapped 90% confidence intervals, can be large for small n but decreases as n increases.

For a fixed value of n, we are interested in the probability that our estimator detects at least half of the discoveries, $\mathbb{P}_{\nu_*}\left(\widehat{\zeta}_n(0) \geq \frac{1}{2}\zeta_*\right)$, which is a function of the fraction of discoveries ζ_* and their effect size γ_*. Our estimator exhibits a sharp transition between detecting fewer than half of the hypotheses and detecting more than half, as shown in Figure 2.5.

We also demonstrate that our method works on distributions other than Gaussians by applying it to synthetic Poisson and binomial data (Figure 2.6). Details of the experiments can be found in Appendix A.4; we note that our test detected the presence of the alternate hypotheses even when no alternates were identifiable via Bonferroni-corrected multiple testing.

2.4.2 Experimental Results on Real Data

We evaluated our estimator on Z-scores from an experiment to identify which genes contribute to influenza replication in *Drosophila*, described by Hao et al. (2008). The data, available in

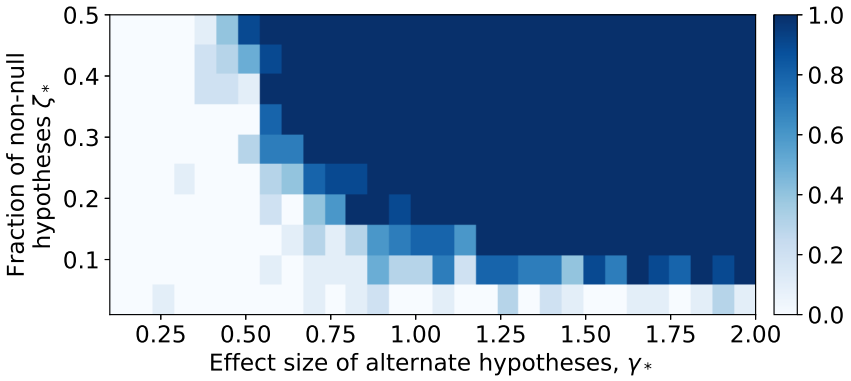


Figure 2.5: Empirical $\mathbb{P}(\widehat{\zeta}_n(0) > \frac{1}{2}\zeta_*)$, for various parameterizations (γ_*, ζ_*) of the two-spike Gaussian model. For a fixed value of $n = 10^4$, the probability of detecting at least half of the discoveries increases as both the effect size γ_* and the fraction of alternate hypotheses ζ_* increase. Empirical probabilities were computed over ten trials.

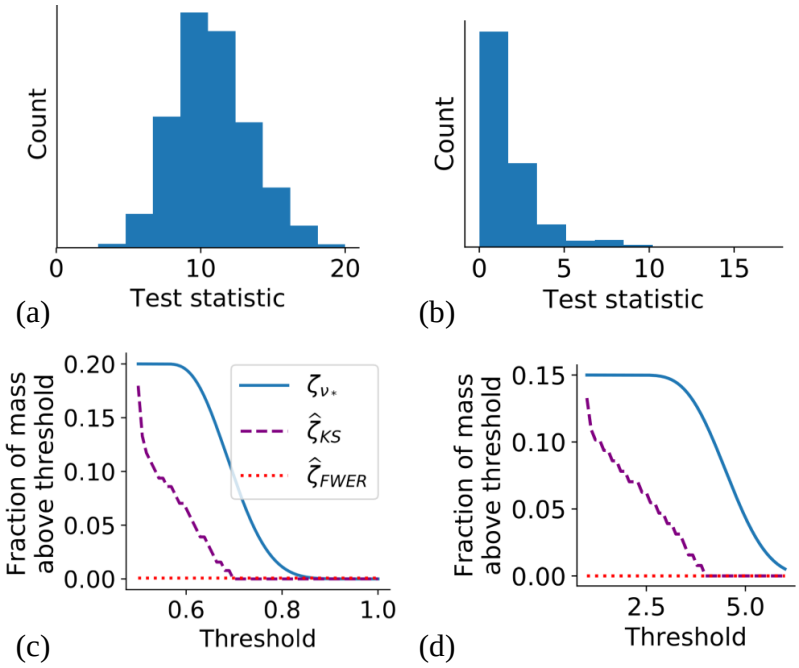


Figure 2.6: Performance of our estimator on non-Gaussian data. (a) and (b) are observations from the binomial and Poisson mixtures, respectively. (c) and (d) show the true fraction of effects above each threshold (blue line), as well as estimates using our method ($\widehat{\zeta}_{KS}$, purple dashes) and identification via Bonferroni-corrected multiple testing ($\widehat{\zeta}_{FWER}$, red dots). Our estimator gets closer to the truth, without overestimating.

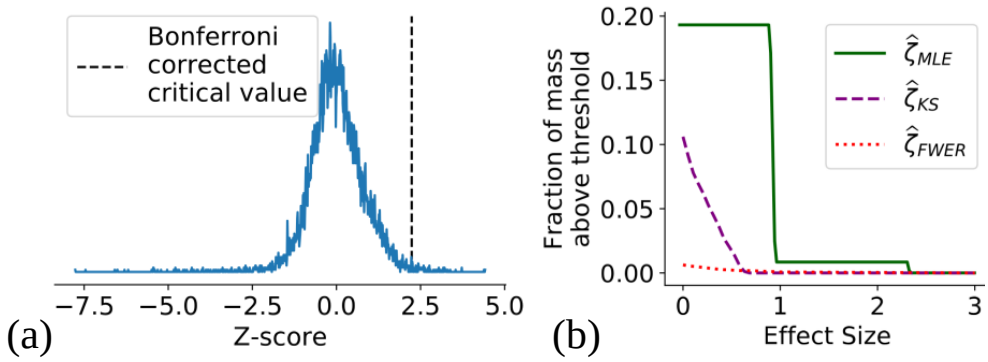


Figure 2.7: Two Z-scores were averaged for each of 13,071 *Drosophila* genes. Even though (a) indicates that very few discoveries could be made, (b) shows that the MLE suggests, and our estimator confirms, many discoveries exist. We note that the MLE provides no guarantee of a conservative estimate, and may drastically overestimate the true fraction at any point.

our supplementary material, consisted of Z-scores from two replicates for each of 13,071 genes. Figure 2.7(a) shows the empirical distribution of the 13,071 averaged Z-scores, which are the observations X_i. The theoretically motivated distribution $X_i \sim \mathcal{N}(\mu, \frac{1}{2})$ is a poor fit to this data, perhaps due to undocumented pre-processing steps not annotated in the dataset, so we began by estimating the variance of these observations. We found that $\sigma^2 = \frac{1}{4}$ provided a good fit to the data; we used this value for the rest of our computations. Testing for significant effect sizes using Bonferroni correction at the 0.05 level (critical value shown in Figure 2.7(a)) resulted in 83 discoveries, representing 0.6% of genes. Given the low number of replicates performed in this experiment, we might suspect that there are more discoveries with smaller effect sizes.

Figure 2.7(b) shows the results of the plug-in MLE estimator $\widehat{\zeta}_{MLE}$, our estimator ($\widehat{\zeta}_{KS}$), and identification with Bonferroni correction ($\widehat{\zeta}_{FWER}$). The fitted MLE suggests that there are around 2600 discoveries to be made (20% of genes), with most effect sizes around 1. As discussed previously, the MLE can overestimate the true number of discoveries and their effect sizes. Our conservative estimator guarantees that there are at least 1400 genes (11% of all genes) with positive effects, including at least 190 genes (1.5%) with effect size of at least 0.5. Our estimator generally detected more discoveries than $\widehat{\zeta}_{FWER}$, excluding the influence of the 23 genes (0.2%) with $X_i > 3$. These observations fall into the *sparse regime* (Section 2.2.2), where our estimator has less power. These results could facilitate the design of an experiment to identify genes with effect sizes exceeding some threshold, or upper bound the cost of a sequential experiment to identify the top 200 genes.

2.5 Discussion and future work

We have presented an algorithm that estimates the fraction of a mixing distribution that lies above some threshold, subject to the constraint that the estimate does not exceed the true fraction with high probability. Our algorithm can be generalized to the following template:

1. Choose some distance metric on CDFs.

2. Find the set $\mathcal{A}$ of "plausible" F_ν given observation $\widehat{F}_n$, which are the CDFs such that $d(F_\nu, \widehat{F}_n) < \tau_{\alpha,n}$.

3. Choose $\tau_{\alpha,n}$ such that $\mathbb{P}(d(\widehat{F}_n, F_{\nu_*}) > \tau_{\alpha,n}) \leq \alpha$.

Returning the minimum amount of mass above the threshold, over the set of plausible distributions $\mathcal{A}$, guarantees with high probability that we do not overestimate the true mass. Our algorithm instantiates this template with the ℓ_∞ norm as the distance metric.

Another natural choice of metric is the likelihood of $\widehat{F}_n$ given F_ν. In order to use this in our template, we need finite sample bounds on the likelihood of $\widehat{F}_n$ given F_{ν_*}. Asymptotic versions of these results are worked out by Jiang & Zhang (2016) for the case of Gaussian X_i, and it would be easy to extend these to finite sample bounds. Extensions of Jiang & Zhang's results should show that the resulting estimator is optimal throughout the sparse and dense regimes. Unfortunately, their value of $\tau_{\alpha,n}$ depends on unknown constants, and therefore it would require extensive simulations with thousands of repetitions for each (ζ, γ) pair to obtain a reliable estimate of the critical values for pilot study analysis. In addition, using the likelihood-based approach for a new distribution f_μ requires an entirely new proof of the high-probability bound on the likelihood ratio.

We believe it would be possible to modify our estimator in order to improve its performance in the sparse regime, where effect sizes are rare but large. Our estimator uses the DKW inequality (Massart, 1990) to measure the plausibility of a latent distribution ν, but the DKW inequality is not tight where the empirical CDF has low variance. Such points occur in the sparse regime, for example at $F_{\nu_*}(\frac{1}{2}\gamma_*)$. Applying a bound that uses variance information, such as an empirical Bernstein DKW as in Lemma 9 of Jamieson and Jain (2018), could address this lack of power.

Any algorithm for this problem necessarily makes some assumptions about the data generating process, otherwise all observations could come from the alternate, with $X \sim P_1$ having density $\widehat{F}_n$. As discussed in Section 2.1.3, previous works have used various assumptions, such as unimodality of ν_* or "purity" of p-values around zero. Our key assumption is the parametric form of X_i, under both the null and the alternate. In practice, in order to decrease our reliance on this assumption, we could learn some parameter of the test statistic distribution from the observations themselves. This was our approach with the *Drosophila* data, when we fit the variance σ^2 of the Z-score. This approach is also taken by Efron (2007). Even more ambitiously, we could learn the conditional distribution $f(X|\mu)$ by fitting it jointly with the means μ, and then use this conditional distribution to generate F_ν from a candidate distribution ν.

Chapter 3

Sample-Efficient Identification of High-Dimensional Antibiotic Synergy with the Normalized Diagonal Sampling Design

Antibiotic resistance is an important public health problem. One potential solution is the development of synergistic antibiotic combinations, in which the combination is more effective than the component drugs. However, experimental progress in this direction is severely limited by the number of samples required to exhaustively test for synergy, which grows exponentially with the number of drugs combined. We introduce a new metric for antibiotic synergy, motivated by the popular Fractional Inhibitory Concentration Index and the Highest Single Agent model. We also propose a new experimental design that samples along all appropriately normalized diagonals in concentration space, and prove that this design identifies all synergies among a set of drugs while only sampling a small fraction of the possible combinations. We applied our method to screen two- through eight-way combinations of eight antibiotics at 10 concentrations each, which requires sampling only 2,560 unique combinations of antibiotic concentrations.

3.1 Introduction

Antibiotic resistance poses a clinical problem for which there are few available solutions. One promising strategy is the use of synergistic antibiotic pairings whose collective potency is greater than expected (Tyers and Wright, 2019). Commercially available examples include the antibiotics trimethoprim and sulfamethoxazole, which inhibit separate steps in the folate biosynthesis pathway (Bushby and Hitchings, 1968), and quinupristin and dalfopristin, which both inhibit the ribosome (Noeske et al., 2014). Very few examples of synergistic combinations exceed two antibiotics (Booth et al., 1994), partly because the number of measurements required to detect multi-antibiotic synergy increases exponentially with the number of antibiotics tested. Exhaustively testing 10 concentrations of five antibiotics would require on the order of 10^5 experiments, which limits the search space even when employing robotics to facilitate experimentation (Tekin et al., 2018). Scaling beyond five antibiotics is therefore impractical, and another approach is needed to explore the space of possible synergies.

Many methods have been developed to search this space. Much of the existing work focuses on identifying combinations with performance exceeding their "expected performance" under some *null model*, i.e., a prior belief about the effectiveness of antibiotics in combination. Measures of antibiotic performance include the absolute *amount* of cells (e.g., optical density after 18 hours)

(Beppler et al., 2016; Katzir et al., 2019), the *rate* at which cells grow (Wood et al., 2012), or the *energy released* by the cells (Kragh et al., 2021). The three most popular null models are Bliss (Bliss, 1939), Highest Single Agent (HSA) (Berenbaum, 1989), and Loewe (Loewe, 1953). Under the Bliss model, antibiotic effects combine probabilistically, so that combining antibiotic A at a concentration that yields a 50% reduction in growth with antibiotic B at a concentration that yields an 80% reduction in growth should result in a 90% reduction in growth (the probability of A *or* B independently killing the bacteria); if the combination has an efficacy greater than 90%, then it is called synergistic. In contrast, the HSA model posits that the effect of a drug combination is equivalent to the maximum of the effects of each drug used by itself. In the preceding example, B is the highest single agent, with an 80% individual reduction in growth, so any combined effect above 80% is considered synergy. Under the Loewe model, antibiotic *doses* combine linearly, so that combining antibiotic A at half of its minimum inhibitory concentration (MIC) and antibiotic B at half of its MIC should result in an inhibitory drug (this is consistent with the "sham combinations" principle, i.e., that half a unit of A plus half a unit of A makes a whole unit of A). If the combination is inhibitory at a lower concentration, it is considered synergistic under the Loewe model. We direct the reader to Foucquier and Guedj (2015) for a comprehensive review of Bliss, Loewe, HSA, and other notions of synergy.

A complementary approach is to define an appropriate combination index (CI) on the space of antibiotic combinations at specific concentrations (Chou and Talalay, 1983; Berenbaum, 1989; Foucquier and Guedj, 2015). The CI of a dose combination measures a deviation from the null model, so that $CI = 1$ implies the combination follows the null model, and $CI < 1$ implies synergy. For *effect-based* null models, such as Bliss or HSA, the CI is naturally defined as the ratio of the prediction of the null model on a combination of drugs to the true effect of that combination. For *dose-effect models*, such as Loewe, the CI is defined relative to underlying isoboles of the same growth level. The Fractional Inhibitory Concentration Index (FICI) is the CI associated with Loewe synergy; an FICI of 1 means the drugs interact according to the Loewe Model (Foucquier and Guedj, 2015). For example, if a combination of drug A at $\frac{1}{4}$ of its MIC and drug B at $\frac{1}{2}$ of its MIC is inhibitory, then the FIC for this combination is $\frac{1}{4} + \frac{1}{2} = \frac{3}{4}$, and the FICI of $A + B$ is therefore no greater than $\frac{3}{4}$. The FICI is interpreted against a standard scale, where values below 1 indicate synergy, while values between $\frac{1}{2}$ and 1 indicate "weak" synergy (Hall et al., 1983). The FICI can also be viewed through a more clinically relevant lens as the minimum fractional concentration among all effective combinations. This perspective leads to a natural optimization problem: find the combination with the lowest normalized dose that is still effective.

Recent research has focused on identifying synergy among more than two drugs. When moving to higher order interactions, a distinction emerges between *total synergy*, which captures the combination's performance relative to its individual components, and *emergent synergy*, which captures the incremental benefit over any subset of the combination; these notions have been defined relative to both the Bliss (Beppler et al., 2016; Tekin et al., 2016; Beppler et al., 2017; Tekin et al., 2018) and Loewe (Cokol et al., 2017) null models. However, a major challenge with identifying higher order interactions is the mounting evidence suggesting that the null models predicted by Loewe and Bliss poorly fit experimental data (Russ and Kishony, 2018; Katzir et al., 2019). In

response, some authors have proposed *data-driven predictive models* trained only on pairs of drugs and then evaluated on three or more drugs, such as the dose model (Zimmer et al., 2016; Katzir et al., 2019), pairs model (Wood et al., 2012; Wood, 2016), and the static λ score (Cokol et al., 2017; Cokol-Cakmak et al., 2020; Yilancioglu and Cokol, 2019). Given that they have no access to data from higher order combinations, predictive models effectively become null models for interactions between more than two drugs. This lets us characterize deviations from the predictive model as "emergent" antagonisms and synergies that arise from higher order effects not predictable by pairs.

Given a specific definition of synergy, a second major challenge is the high measurement burden of identifying synergistic combinations. For example, 10^8 measurements are required to exhaustively test eight drugs at 10 concentrations each. Strategies to overcome exhaustive sampling fall into two categories: *parametric modeling* and *experimental design*. The former approach applies concepts from machine learning to build parametric models that can potentially predict all possible drug combinations accurately but can be learned using only a fixed number of parameters. Models of this form include the previously mentioned dose (Zimmer et al., 2016) and pairs (Wood, 2016) models, which explicitly assume no higher order interactions beyond pairs of drugs; mechanistic models, which use knowledge of the underlying drug targets (Kavčič et al., 2021) or gene expression data (Lukačišin and Bollenbach, 2019); and the MAGENTA model (Cokol et al., 2018), which leverages phenotypic information about the cell's response to antibiotics.

The experimental design approach is a complementary strategy that reduces the amount of data to be collected, sometimes at the cost of data fidelity. For example, some studies severely restrict the number of doses per antibiotic combination (Tekin et al., 2018) in order to exhaustively sample the space of possible combinations. Other studies employ *diagonal sampling*, where the relative proportions of the antibiotics remain fixed but the absolute quantities vary (Berenbaum, 1977; Cokol et al., 2017; Katzir et al., 2019). The diagonal sampling method has been proposed as a way to feasibly sample in higher dimensions and has been justified with the claim that the diagonal "provides the most information about the shape of the contour [phenotype isobole]" (Cokol et al., 2017). While the validity of diagonal sampling for Loewe synergy has received experimental support in some studies (Yilancioglu and Cokol, 2019; Cokol-Cakmak et al., 2020), to date no work has rigorously justified its use or provided any guarantee about what kinds of synergies a diagonal design may or may not uncover. Absent such a rigorous justification, any study that fails to find synergy leaves open the possibility that synergy may still exist.

This chapter describes three major contributions: (1) we develop a novel combination index, the Minimax Effective Concentration Index (MECI), which naturally extends the FICI and HSA models; (2) we present a theoretical framework for determining total and emergent synergies based on the MECI; and (3) we propose a new experimental design for provably identifying the MECI, called *normalized diagonal sampling* (NDS). The MECI has dual advantages over previously proposed metrics. It can be efficiently identified in high dimensions using the NDS design; under realistic assumptions about the behavior of antibiotics in combination, we prove that the NDS design finds the MECI with exponentially fewer samples than the full factorial experiment that samples all drugs at all combinations of concentrations. Finding the MECI of an 8-drug combination using 10 concentrations requires only $10 \cdot 2^8 \approx 2500$ samples, whereas finding the FICI in the

same setting via the full factorial design would require an infeasible 10^8 samples. The second advantage of the MECI is the flexibility it gives the experimentalist to capture clinically relevant drug combinations. The definition of the MECI includes a normalization factor for each drug, and the choice of normalization can be tailored to the goals of the synergy study at hand. We present two examples of the normalization factor in Section 3.2.3.

3.2 Methods

3.2.1 The Minimax Effective Concentration Index

In clinical practice the goal is to administer antibiotic combinations that are effective while avoiding high doses, which may cause adverse effects. Antibiotic doses are measured relative to various metrics, such as the strain-specific MIC or species-specific breakpoints, determined by bodies such as the European Committee on Antimicrobial Susceptibility Testing (EUCAST) (2021). In this section, we define the *Minimax Effective Concentration Index (MECI)*, which captures the idea of avoiding high doses by minimizing the highest single antibiotic's concentration (appropriately normalized) among antibiotic combinations that are "effective" at inhibiting growth. The precise definition of an "effective combination" is specified by the practitioner but could, for example, represent whether or not growth at a predefined time point remains below a pre-specified yield threshold.

Given a definition of an effective drug combination, we can now define the MECI. Let $\Omega = \{1, 2, 3, \ldots, d\}$ index a set of d drugs, where drug i will be tested at $m-1$ concentrations $\chi_i = \{x_{i,1}, x_{i,2}, \ldots, x_{i,m-1}\}$. The concentrations of drug i are normalized to concentration N_i for analysis purposes so all drugs can be compared on a similar scale. In our experiments, we chose to test each drug at the same set of normalized ratios $\frac{x_{i,j}}{N_i}$.

Define $\mathcal{X}(\Omega) := \prod_{i\in\Omega}(\chi_i \cup 0)$, the set of all possible combinations of all subsets of the d drugs in Ω at their $m-1$ concentrations. A single $\mathbf{x} \in \mathcal{X}(\Omega)$ represents a single experimental condition, with x_i encoding the concentration of drug i (or $x_i = 0$ if drug i is absent). The set $\mathcal{X}(\Omega)$ represents all m^d experimental conditions that would be tested in the full-factorial experimental design.

We define the *minimax concentration* of any $\mathbf{x} \in \mathcal{X}(\Omega)$ as the maximum normalized concentration among all drugs in the set Ω, i.e., $\max_{i\in\Omega} \frac{x_i}{N_i}$. The *Minimax Effective Concentration Index (MECI)* is the smallest minimax concentration among all effective combinations $\mathbf{x} \in \mathcal{X}(\Omega)$. Specifically, the MECI of a set of drugs Ω can be expressed succinctly as the solution to an optimization problem:

$$MECI(\Omega) = \min_{\mathbf{x}\in\mathcal{X}(\Omega)} \max_{i\in\Omega} \frac{x_i}{N_i} \quad \text{such that} \quad \mathbf{x} \text{ is effective} \tag{3.1}$$

The MECI($\{i\}$) of a single antibiotic i is the normalized concentration at which i is individually effective. For example, if N_i is the MIC of drug i and a drug is considered "effective" if it completely inhibits growth, then MECI($\{i\}$) $= 1$.

When screening for synergy, it is important to identify combinations where each antibiotic's inclusion is justified by the effectiveness it brings to the combination. Indeed, synergies among lower order antibiotic combinations are likely to be detected in higher order combinations that include those same drugs; therefore, finding clinically useful combinations requires determining whether all antibiotics in a combination are necessary to achieve a synergistic effect. To this end, we develop two scores for interpreting whether a drug combination provides a meaningful improvement over its components: the Total Synergy Score and the Emergent Synergy Score. The *Total Synergy Score (TSS)* measures the reduction in maximum concentration attainable by the combination as compared to each individual drug. For any combination of drugs Ω, it can be expressed mathematically as

$$TSS(\Omega) = \frac{MECI(\Omega)}{\min_{i\in\Omega} MECI(\{i\})}. \tag{3.2}$$

If the normalization N_i is the MIC of each drug, then the denominator $\min_{i\in\Omega} MECI(\{i\}) = 1$, and the TSS is identical to the MECI. However, if the normalization can potentially be very far from the MIC, reporting the TSS ensures that a combination must improve upon its best single ingredient in order to be considered synergistic.

We further define the *Emergent Synergy Score (ESS)* to capture the improvement offered by the combination when compared to any of its subsets, following the intuition that the combination must significantly improve upon any subset of its ingredients to justify the use of additional drugs. The ESS is defined as

$$ESS(\Omega) = \frac{MECI(\Omega)}{\min_{\Omega'\subset\Omega} MECI(\Omega')}. \tag{3.3}$$

We observe that $TSS(\Omega) \leq ESS(\Omega) \leq 1$. An ESS score of 1 indicates that some proper subset of Ω has an MECI at least as small as Ω itself; this can occur in instances of indifference or antagonism. The TSS and ESS scores identify only synergy, not antagonism; for a discussion of antagonism, see Section 3.4.

While any ESS less than 1 indicates emergent synergy, we suggest as interpretive criteria that conclusions of *synergy* be restricted to combinations with an ESS of 0.25 or below, while combinations with an ESS between 0.25 and 1 should be interpreted as evidence of *weak synergy*. These interpretive criteria are consistent with the FICI interpretation of weak synergy when two drugs are tested according to the NDS design, since the cutoff FICI of 0.5 and the cutoff ESS of 0.25 both occur when two drugs are combined at $\frac{1}{4}$ of their MICs to yield an effective combination. Figure 3.1 illustrates the calculation of the MECI, TSS and ESS for a three-drug experiment.

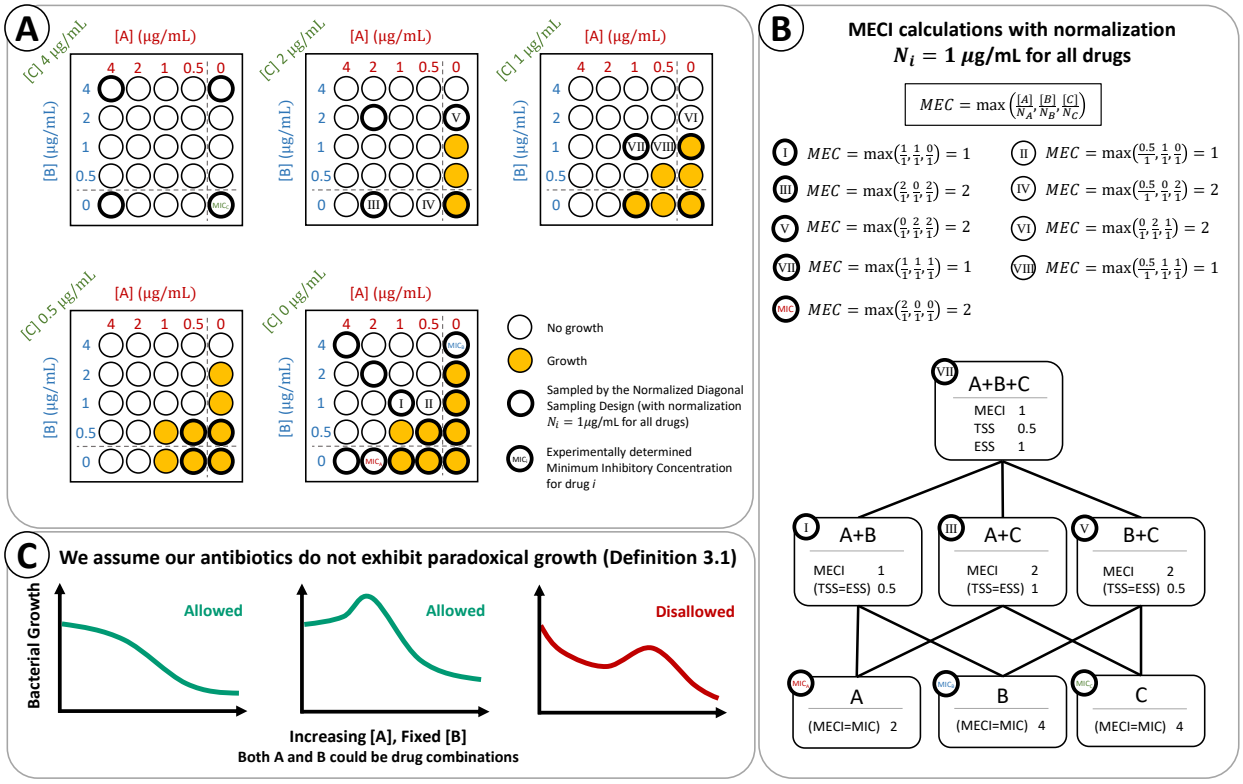


Figure 3.1: Calculation of the MECI, TSS, and ESS for a three-drug experiment. (A) Testing three drugs at four concentrations each could be performed exhaustively using a three-dimensional checkerboard assay, as depicted here. When the effectiveness measure is taken to be the absence of visible growth, the MEC can be calculated for each effective measurement in the checkerboard assay. (B) Computation of select MEC values from the checkerboard assay. The MECI of a drug combination is the minimum MEC among tested combinations. The MECI, TSS, and ESS of all subsets are computed according to their definitions, with normalization $N_i = 1$ μg/mL for all drugs. The well in the upper-left corner of each combination's calculated values witnesses the MECI of that combination; we see that the wells tested by the NDS design (shown with bold edges) are sufficient to identify the MECI. (C) If the behavior of the antibiotics satisfies Definition 3.3.1, we prove that the NDS design always identifies the MECI for every combination tested. We illustrate some examples of allowed and disallowed behavior of the measured response as a function of increasing some antibiotic (combination) A as the concentration of antibiotic (combination) B remains fixed.

3.2.2 Relationship to other metrics

For comparison to an established metric, we note that the FICI, like the MECI, can be written as the solution to an optimization problem. The FICI of combination Ω is given by

$$FICI(\Omega) = \min_{\mathbf{x} \in \mathcal{X}(\Omega)} \left(\sum_{i \in \Omega} \frac{x_i}{MIC_i} \right) \quad \text{such that} \quad \mathbf{x} \text{ is effective.} \tag{3.4}$$

Note that the FICI differs from the MECI in two important ways. First, the FICI measures the amount of drug in each combination using its summed fractional concentration $\sum_{i \in \Omega} \frac{x_i}{MIC_i}$, whereas the MECI uses the maximum fractional concentration $\max_{i \in \Omega} \frac{x_i}{N_i}$. Second, the FICI exclusively normalizes to the MIC of each drug, while the MECI allows flexibility in the choice of normalization.

We conclude by observing that the MECI corresponds to a dose-effect interpretation of the HSA model. Intuitively, the MECI minimizes the minimax concentration among effective combinations, where the minimax concentration is exactly the highest (normalized) concentration of a single agent. This connection is made precise in Appendix B.2.

3.2.3 Choice of normalization

The MECI provides flexibility in the choice of normalizing metric N_i, which allows the experimentalist to tailor the choice of N_i to the goals of the investigation. In particular, different choices of N_i lead to different interpretations of synergy. For example, a natural choice of normalization N_i is the minimum inhibitory concentration for each drug i. With this choice, a low ESS or TSS indicates that the combination is effective at concentrations far below the individual drugs' MICs, which aligns with the Loewe (and, to a lesser extent, Bliss) conceptions of synergy.

Another meaningful choice of N_i could be a strain-independent metric like the EUCAST breakpoint, which is the highest concentration at which the organism is considered sensitive to the antibiotic (2021). For example, when testing for synergy in a multi-drug-resistant strain, a clinically important goal could be to find a combination in which each drug is present below its breakpoint concentration, so that $MECI \leq 1$.

We observe that it is possible for a drug combination to be synergistic with respect to one normalization N_i but not another, or for the strength of the interaction to depend upon the choice of normalization. As a result, it is important for the choice of N_i to reflect the scientific or clinical goal of the experiment and for this choice to be reported along with the synergy scores.

3.2.4 The normalized diagonal sampling design

We introduce the *normalized diagonal sampling (NDS) design*, an experimental design that samples all combinations of drugs in Ω at equal concentrations relative to the normalizations N_i. We begin with a set of ratios at which each drug will be tested, $\frac{x_i}{N_i} \in \{c_1, c_2, \ldots, c_{m-1}\}$. For example, the ratios may be powers of two, $c_j = 2^{-j}$, in which case the drugs are tested on a two-fold dilution gradient. The spacing and number of concentrations should be chosen to balance the

resolution of the synergies detected with the number of experiments required, while testing a sufficient range of concentrations so that all drug combinations can be observed at both *effective* and *ineffective* concentrations. For all subsets $\Omega' \subseteq \Omega$, the NDS design evaluates the combination Ω' at all $m-1$ normalized concentrations $\{c_1\mathbf{1}, c_2\mathbf{1}, \ldots, c_{m-1}\mathbf{1}\}$, where $\mathbf{1}$ represents a vector of all ones of size $|\Omega'|$. In Section 3.3.1 we prove that, under mild assumptions about the behavior of antibiotic combinations, normalized diagonal sampling provides strong theoretical guarantees on the detection of synergy. In particular, the design provably identifies the MECI from a collection Ω of d antibiotics without having to sample all $|\mathcal{X}(\Omega)| = m^d$ possible combinations. Since the NDS design provably finds the MECI, high-dimensional antibiotic combination screens can be run with the confidence that if no synergies are identified, then none exist.

Exhaustive tests for synergy are conducted with checkerboard assays (see Figure 3.1) requiring m^d wells to screen d drugs at m concentrations each. The NDS design significantly reduces the required number of wells by testing along the "diagonal" — testing each combination with all drugs present at the highest concentration, then at the second-highest concentration, and so on. Under the NDS design, each of the 2^d drug combinations requires only m wells, for a total requirement of $m \cdot 2^d$ wells. For eight drugs and 10 concentrations per drug, this requires $m \cdot 2^d = 10 \cdot 2^8 = 2,560$ wells, or about twenty six 96-well plates. This is experimentally feasible, whereas $m^d = 10^8$ wells (requiring approximately 10^6 plates) is not.

3.2.5 Quantification of bacterial growth

Computing the MECI requires the analyst to specify an *effectiveness metric* that specifies whether the antibiotics tested were *effective* or *ineffective* at the given concentrations $\mathbf{x}$. For our experiments, we considered an antibiotic combination *effective* in inhibiting bacterial growth if the area under the growth curve (AUGC) was less than a predefined threshold. The AUGC measures the area under the curve of optical density (at 600nm) over time after subtracting the background OD reading. The area was approximated using a trapezoidal Riemann sum, taken in 15 minute increments over the 24 hours following inoculation. We also considered quantifying effectiveness using the maximum growth rate, computed as the maximum slope of a five-point moving average of the log optical density versus time curve; since this technique agreed with the AUGC, we did not include it in our results. Any measurement that captures the notion of antibiotic effectiveness could be used to compute the MECI.

3.2.6 Experimental conditions

Experiments were performed using the wild-type *E. coli* strain MG-1655 (NR-2653; BEI Resources, Manasses, VA, USA) in BBL Mueller Hinton II Broth (Cation-Adjusted) (BD Diagnostics, Spark, MD, USA). All experiments were performed in duplicate in 96 well plates and were fully randomized across well, plate and day of experimentation with the use of the OT-2 liquid handling robot (Opentrons, Brooklyn, NY, USA). In each well, antibiotics were diluted into $200\ \mu L$ of media, then $50\ \mu L$ of a 10^{-4} dilution of *E. coli* overnight culture, incubated at 37°C in BBL Mueller Hinton II Broth (Cation-Adjusted), was added. A fresh preparation of overnight culture was used for each

day of experimentation. Plates were sealed with a gas-permeable sealing membrane (Breathe-Easy, Sigma-Aldrich, St. Louis, MO, USA) and incubated at 37°C for 24 hours, during which time optical density readings (600nm) were taken at 15 minute increments using a Biotek BioStack II coupled to a Biotek Epoch II Microplate Spectrophotometer. Plates were orbitally shaken for 15 seconds prior to each reading.

Eight antibiotics were chosen for their diversity of class and mechanism of action: ampicillin, aztreonam, ceftazidime, chloramphenicol, ciprofloxacin, gentamicin, trimethoprim, and tobramycin. Antibiotics were dissolved in water with the exception of aztreonam, ceftazidime and trimethoprim, which were dissolved in DMSO, and chloramphenicol, which was dissolved in ethanol. Table 3.1 identifies the classes of each antibiotic used, the EUCAST susceptible breakpoints (2021), and the minimum inhibitory concentrations determined experimentally.

3.3 Results

The full-factorial sampling design is intractable beyond a small number of antibiotics, which motivates the need for a more efficient design. We begin with a proof of correctness for the normalized diagonal sampling design, showing that the NDS design identifies the MECI using significantly fewer measurements than the full factorial design. The NDS design samples along the diagonal in multi-antibiotic concentration space, letting us find all synergies among combinations of up to eight antibiotics at 10 concentrations with only 2,560 samples. This scale of experiment can be effectively pipetted by hand or using robotics, as we did here. Using robotics enabled us to perform complete pipetting randomization, minimizing plate position effects. In combination, the NDS design and robotics let us confidently identify synergies while mitigating the influence of experimental artifacts that complicate large-scale experiments.

3.3.1 The NDS design provably finds the MECI

We now show that the NDS design provably identifies the MECI of a set of drugs using significantly fewer samples than the full factorial design. Since the NDS design for a set of drugs Ω also involves performing the NDS design for all sets $\Omega' \subset \Omega$, it correctly identifies the MECI, TSS and ESS for each $\Omega' \subseteq \Omega$.

The proof begins with an assumption that adding an antibiotic A to a fixed concentration of antibiotic B does not exhibit *paradoxical growth*, that is, once increasing the amount of A reduces the level of the measured response (e.g., growth), further increasing A cannot increase the response. Next, we observe that a fixed-ratio combination of drugs itself behaves like an antibiotic, with its own dose response curve and its own interactions with other (combination) antibiotics. This perspective lets us extend the assumption to the general case where A and B are themselves antibiotic combinations, stated formally in Definition 3.3.1. When the effectiveness measure in the definition of MECI behaves according to Definition 3.3.1, we can declare many combinations ineffective without ever measuring them because they lie between points already measured to be ineffective. Without such an assumption, we would have no way to know whether a point x is

effective without testing it. The formal theorem statement is provided in Theorem 3.3.2; the proof is available in B.1.

Definition 3.3.1 (Absence of paradoxical growth)**.** *Let Ω be a set of antibiotics. Let the vector $\mathbf{x_0} \in \mathbb{R}_{\geq 0}^{|\Omega|}$ represent a fixed background concentration of antibiotics to which we add increasing amounts of another antibiotic combination $\mathbf{x} \in \mathbb{R}_{\geq 0}^{|\Omega|}$. We say the set of drugs Ω* does not exhibit paradoxical growth *if, for all $c_3 > c_2 > c_1 \geq 0$, the response $r : \mathbb{R}_{\geq 0}^{|\Omega|} \to \mathbb{R}$ satisfies*

$$r(\mathbf{x_0} + c_2\mathbf{x}) < r(\mathbf{x_0} + c_1\mathbf{x}) \implies r(\mathbf{x_0} + c_3\mathbf{x}) \leq r(\mathbf{x_0} + c_2\mathbf{x}). \tag{3.5}$$

Figure 3.1c illustrates our assumption by showing several allowed and disallowed shapes of the dose-response curve as some combination A is added to the base combination B. We specifically note that this assumption does not preclude so-called "hyper-antagonism," where the addition of A to B yields a less effective response than B alone. Such behavior is allowed under our assumption as long as the following condition holds: once increasing concentration of A starts reducing the response r, further increasing the concentration continues to reduce the response.

How plausible is the assumption of non-paradoxical growth? When antibiotic effectiveness is measured in a broth dilution assay, as in our experiments, we are aware of no published evidence of paradoxical growth. When effectiveness is measured using a survival assay, in which bacteria are first treated with antibiotics for a specified time and then the culture is grown in the absence of antibiotics, paradoxical growth is known as the *Eagle effect*, first observed by Eagle and Musselman (1948) and reviewed by Prasetyoputri et al. (2019). If any antibiotic or combination in an experiment displayed the Eagle effect, then the NDS design could fail to identify synergies if drug effectiveness were quantified using a survival assay. To further support our assumption of non-paradoxical growth, we tested 100 randomly chosen ratios of antibiotics against fixed background combinations of antibiotics, the precise setting in Definition 3.3.1. Appendix B.3 shows the results of these experiments, in which no paradoxical growth was observed.

Suppose we define a combination of drugs at a given concentration $\mathbf{x}$ as *effective* whenever the response falls below some threshold ($r(\mathbf{x}) \leq t$). Then, as long as the response behaves according to Definition 3.3.1, we can identify entire regions of the antibiotic combination space as *ineffective* using only measurements on the boundary of the space. This leads to our main result: the correctness of the NDS design in the absence of paradoxical growth.

Theorem 3.3.2. *Assume the set of drugs Ω does not exhibit paradoxical growth (Definition 3.3.1). Then, the normalized diagonal sampling design applied to Ω identifies $MECI(\Omega')$ for all $\Omega' \subseteq \Omega$.*

The proof of Theorem 3.3.2 is deferred to Appendix B.1. We observe that the NDS design provably identifies $MECI(\Omega)$ with significantly fewer samples than the naive full factorial design: the full factorial design requires m^d samples, while the NDS design requires only $m \cdot 2^d$.

Table 3.1: The eight antibiotics chosen for our studies, including EUCAST susceptible breakpoints (2021) and experimentally determined MICs for *E. coli* strain MG-1655.

Drug (ASM Code)	ASM Code	Product Number	Antibiotic Class	EUCAST Susceptible Breakpoint (μg/mL)	MIC (μg/mL)
Ampicillin	AMP	BP1760 (Fisher Scientific)	β-lactam (penicillin)	8	16
Aztreonam	ATM	15151 (Chem-Impex)	β-lactam (monobactam)	1	0.25
Ceftazidime	CAZ	AC461730050 (Acros Organics)	β-lactam (cephalosporin)	1	0.25
Chloramphenicol	CHL	C0378 (Sigma-Aldrich)	Amphenicol	8	8
Ciprofloxacin	CIP	199020 (MP Biomedicals)	Quinolone	0.25	0.015625
Gentamicin	GEN	00149 (Chem-Impex)	Aminoglycoside	2	0.5
Trimethoprim	TMP	92131 (Sigma-Aldrich)	Antifolate	4	0.25
Tobramycin	TOB	455430010 (Acros Organics)	Aminoglycoside	2	0.5

3.3.2 Experimental results

Identification of emergent effects relative to breakpoint

Using the NDS design, we performed experiments to screen all 2^8 combinations of the 8 drugs shown in Table 3.1. The set of drugs was chosen to cover a wide range of mechanisms among drugs with defined EUCAST breakpoints for *Enterobacterales*.

Effectiveness was specified as complete inhibition of growth, measured by AUGC (as described in Section 3.2.5). The normalization constant N_i for the first experiment, used for both the NDS design and for calculating the TSS and ESS, was the EUCAST susceptible breakpoint (2021) for *Enterobacterales*. As motivated in Section 3.2.3, normalization relative to the breakpoint provides a strain-independent measure of antibiotic interactions, which may be more relevant to how drugs are prescribed in combination clinically. We tested concentrations across 10 steps of a two-fold dilution gradient so that any interactions could be identified within a power of two.

We computed MECI, ESS and TSS for each combination of two through eight antibiotics. Histograms of the ESS and TSS scores across all combinations, plotted by the number of drugs combined, are shown in Figure 3.2a (recall that a lower ESS score indicates greater synergy, and a $\log_2$ ESS score of 0 indicates the absence of synergy). Among all 2^8 combinations tested, no combination met our synergy threshold of a $\log_2$ ESS of -2 or lower (4-fold synergies), although several combinations exhibited weak synergy with a $\log_2$ ESS of -1. This indicates only weak synergy among all combinations of these eight antibiotics, consistent with previous studies that found synergy to be rare (Cokol-Cakmak et al., 2020; Tekin et al., 2018; Chandrasekaran et al., 2016; Yilancioglu and Cokol, 2019). As anticipated, the most commonly reported $\log_2$ emergent effect was zero. Figure 3.2b shows the makeup of the combinations exhibiting weak synergy. Full data, including the ESS and TSS of each of the 2^8 combinations and a Loewe analysis showing the absence of strong Loewe synergy along the diagonals tested, is available in Appendix B.4.

We emphasize that our results come with strong guarantees under the assumption that the drugs do not exhibit paradoxical growth. Since our experiments were conducted according to the NDS design, Theorem 3.3.2 *guarantees* that, among these eight drugs, combination has a $\log_2$ ESS

score of -2 or less when ESS measures synergy relative to the EUCAST breakpoint. In particular, the result guarantees that performing the full factorial experiment would result in the same ESS and TSS scores that we found with the NDS design. If the goal is to minimize the maximum amount of drug applied relative to the EUCAST breakpoint, we therefore conclude that no significant gains are possible by combining this set of drugs against the strain we tested and under our experimental conditions.

Identification of emergent effects relative to MIC

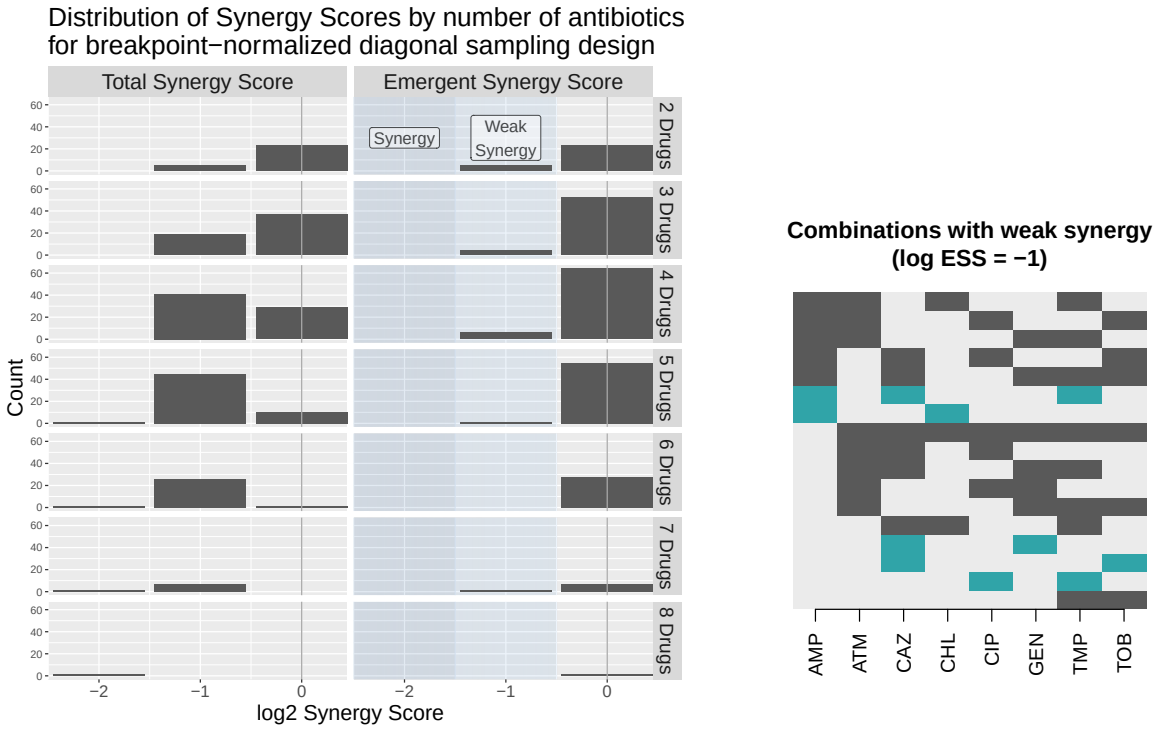
When compared to earlier methods such as the FICI, an important advantage of our synergy screening method is the flexibility provided by the choice of the normalizing constant N_i. To demonstrate the utility of normalization, we repeated the preceding experiment but with all concentrations normalized to the MIC of the individual drugs (listed in Table 3.1). Normalizing to the MIC means that all drugs are combined at similar points along their dose-response curves, which we might expect to result in larger interactions. It also has the benefit of mirroring the FICI synergy screen, where the "fractional concentration" of a combination is always measured relative to the MICs of the individual antibiotics. For this reason, the MIC-normalized experiment is more directly comparable to previous synergy screening techniques.

We computed the MECI, TSS and ESS for each combination of antibiotics and show the results in Figure 3.3. We see that the most common $\log_2$ ESS score is again 0, and that the lowest observed $\log_2$ ESS score is -1. The lack of strong synergy remains consistent with previous literature showing that synergy is rare (Cokol-Cakmak et al., 2020; Tekin et al., 2018; Chandrasekaran et al., 2016; Yilancioglu and Cokol, 2019). Again, we emphasize that the NDS design provides a strong guarantee in the absence of paradoxical growth: not only did we not find any combinations with $\log_2$ ESS of -2 or lower, we also know that there is no ratio at which the drugs can be combined that will exhibit an ESS of this magnitude. Figure 3.3b shows the combinations exhibiting weak synergy.

Compared to the breakpoint-normalized experiment, we observe that the results of the MIC-normalized experiment show more weak synergies and more negative $\log_2$ TSS scores. One possible explanation is that the breakpoint-normalized experiment combined drugs at concentrations with very different individual effectiveness. For example, ciprofloxacin has a breakpoint that is four steps above its MIC for this strain. As a result, combinations of ciprofloxacin and any other drug at concentrations several steps below the breakpoint may behave like ciprofloxacin alone, resulting in a $\log_2$ TSS of zero.

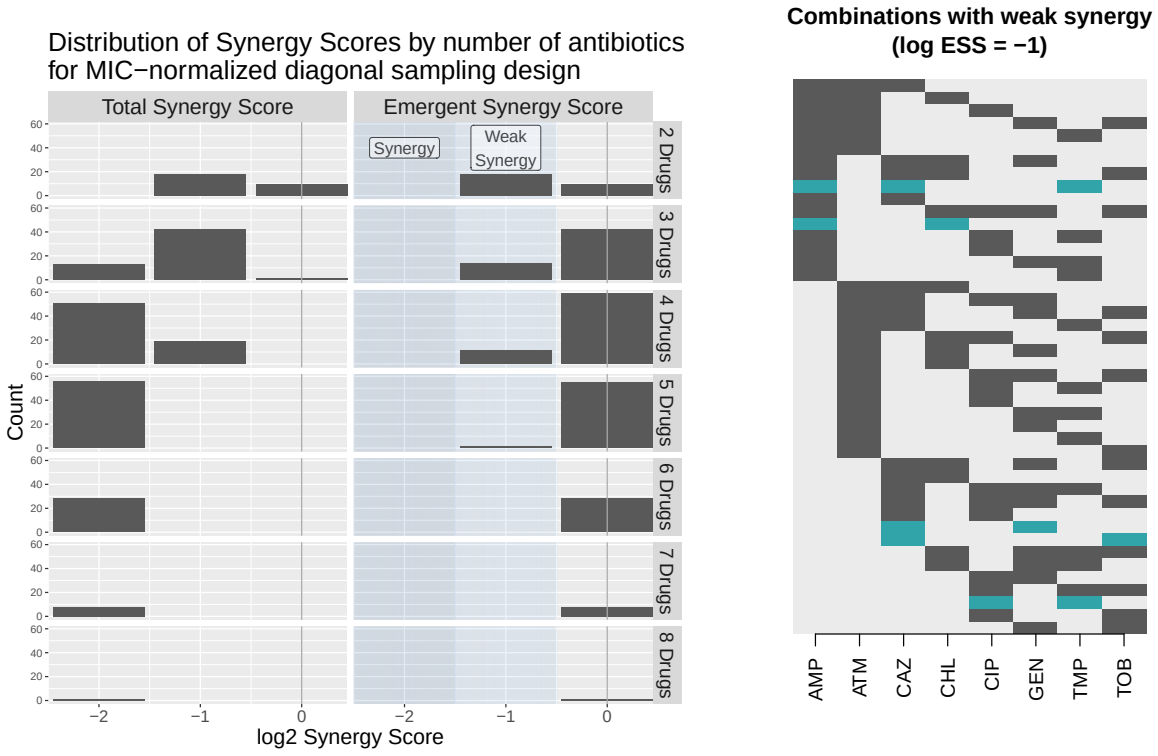
3.4 Discussion

We developed a novel scoring function to identify antibiotic interactions, the Minimax Effective Concentration Index, which applies the principles of the HSA model to determine whether an antibiotic combination is synergistic. We extend the highest single agent model into higher dimensions from both a theoretical perspective, developing a framework for identifying total and emergent synergies, and from a practical perspective, introducing a novel experimental design



(a) (b)

Figure 3.2: Results from the breakpoint-normalized experiment. (a) Distribution of TSS and ESS scores across all 2^8 different breakpoint-normalized combinations of the eight antibiotics in Table 3.1, separated by the number of drugs in the combination. (b) Representations of the 17 drug combinations exhibiting weak synergy. Each row represents one combination; dark shades (black and blue) indicate presence of the drug, while light gray indicates absence. Black represents combinations that are weakly synergistic according to the breakpoint normalization but not the MIC normalization (next section), while blue shows the five combinations that exhibited weak synergy according to both normalizations.



(a) (b)

Figure 3.3: Results of the MIC-normalized experiment. (a) Distribution of TSS and ESS scores across all 2^8 different MIC-normalized combinations of the eight antibiotics in Table 3.1, separated by the number of drugs in the combination. (b) Representations of the 44 drug combinations exhibiting weak synergy; combinations that were also weakly synergistic under the breakpoint normalization are shown in blue. Each row represents one combination; black/blue indicates presence of the drug, while gray indicates absence.

that provably identifies all total and emergent synergies. Applying our methodology to study combinations of eight antibiotics yielded no clinically relevant synergies, which we define as emergent synergies with a 4-fold decrease or more in concentration when compared to their best subset. The strong theoretical guarantees of our sampling scheme allow us to conclude that there is no clinically relevant synergy among these eight drugs under the experimental conditions and normalization methods we considered, even though we sampled only a small fraction of the possible space of concentrations.

The MECI also has clinical relevance aside from the classification of synergistic combinations. When drug combinations are prescribed in clinical practice, each individual drug is typically administered at its standard dose. If the MECI is normalized relative to that standard dose, then an MECI < 1 indicates that the drugs can each be administered at doses less than their standard dose while still remaining effective. In addition, if NDS is used to identify the MECI and paradoxical growth is absent, then there is no effective combination at doses less than the experimentally determined MECI.

We emphasize that the theoretical guarantees of our method apply only when the set of antibiotics exhibits non-paradoxical growth, which we define as the display of unimodal dose-response curves along any constant-ratio drug combination in the presence of any fixed background drug combination. Our assumption of non-paradoxical growth may be violated when the antibiotics exhibit the Eagle effect and the response is measured using a survival assay, or if higher order combinations of drugs behave very differently from our intuition based on the one- and two-drug settings. We further emphasize that these guarantees only apply to synergy as determined by the MECI, and that our design is not guaranteed to identify Loewe or Bliss synergies.

While the NDS design provably identifies all synergies relative to a given normalization, it is limited in that it cannot provably identify all antagonisms. Recall that the MECI quantifies the smallest minimax concentration among drug combinations measured to be effective. Another perspective on this optimization problem is to consider multiple rays representing different concentration ratios (see Foucquier and Guedj (2015), Figure 4a), which could be considered different "diagonals," in the language of diagonal sampling. The MECI minimizes the minimax effective concentration along each of those rays. An analogous definition of *antagonism* from this perspective would be a new combination index that *maximizes* the minimax concentration of *ineffective* combinations among these rays. Total and emergent antagonism could then be defined analogously to their definitions for synergy. Unfortunately, the NDS design's theoretical guarantees extend only to minimizing the minimax effective concentration (identifying the greatest synergy), not to maximizing the minimax ineffective concentration (identifying the greatest antagonism). The difference is due to an asymmetry in Definition 3.3.1; it is possible for the response to increase and then decrease as antibiotic B is added to a fixed concentration of A (so-called *hyperantagonism*), but it is not possible for the response to decrease and then increase (paradoxical growth). Consequently, while an ESS of 1 indicates a lack of synergy, an ESS of less than 1 does not rule out the existence of antagonism in some part of the antibiotic combination space. Notwithstanding this limitation, we believe synergies are typically the most clinically relevant, so the inability to find antagonisms is therefore not too limiting in practice.

The NDS design makes provably identifying high-order interactions experimentally feasible for the first time. With this new method, it is now possible to screen for interactions among more antibiotics, across additional strains, and on multiple media, providing information on whether interactions are conserved across these variables. The NDS design is simple to implement: given a definition of antibiotic normalization (e.g., MIC), it samples along the diagonals in multi-antibiotic space. We conclude by noting that the interaction among biologically active agents is not only of interest for antibiotics, but also for many other biologically active agents, such as immunosuppressants (Berenbaum, 1977), environmental toxins (Cedergreen, 2014), anesthetics (Hendrickx et al., 2008), and anticancer drugs (Scripture and Figg, 2006; O'Neil et al., 2016; Kurtz et al., 2018; Zimmer et al., 2016). Whenever the absence of paradoxical growth can be assumed, the NDS design can be applied to identify synergy in high dimensions.

Chapter 4

Analysis and Methods to Mitigate Effects of Under-Reporting in Count Data

Under-reporting of count data poses a major roadblock for prediction and inference. In this chapter, we focus on the Pogit model, which deconvolves the generating Poisson process from the censuring process controlling under-reporting using a generalized linear modeling framework. We highlight the limitations of the Pogit model and address them by adding constraints to the estimation framework. We also develop uncertainty quantification techniques that are robust to model mis-specification. Our approach is evaluated using synthetic data and applied to real healthcare datasets, where we treat in-patient data as 'reported' counts and use held-out total injuries to validate the results. The methods make it possible to separate the Poisson process from the under-reporting process, given sufficient expert information.

4.1 Introduction

Under-reporting of count data is a pervasive problem in many fields, including econometrics (Winkelmann, 1996), epidemiology (Stoner et al., 2019), and engineering (Wood et al., 2016). In population health, under-reporting of key statistics such as injuries and birth defects impedes estimation of the burden of disease (Vos et al., 2020). The goal of statistical modeling in this area is to develop a rigorous approach to deconvolve the data-generating from the under-reporting mechanism. However, accurately separating these two mechanisms is extremely challenging without contextual information, such as covariates related to under-reporting and those related to the true data-generating process.

The predominant approach to modeling under-reported counts focuses on a two-stage model. Events are first generated according to a Poisson process with mean λ. They are then reported according to a binomial process with probability p, with resulting reported events having mean $\mu = \lambda p$. If covariate information is available, both λ and p can be modeled as functions of those covariates; otherwise, a mathematically convenient distribution over these parameters is assumed. When the probability of reporting is modeled using logistic regression (i.e., with a logit transform), the resulting model is called a *Poisson-Logit (Pogit) model.*

Unfortunately, the deconvolution problem to separate the Poisson and bionomal processes, i.e., inferring λ and p from data that inform μ, is very difficult. Since only reported counts are observed, it is challenging to determine whether these counts result from many events with a low rate of reporting (large λ, small p) or few events with a high rate of reporting (small λ, large

p). This problem is exacerbated when covariates are shared between the Poisson and binomial processes or when they are highly correlated, and results in a wide range of plausible solutions. Addressing this variability, in both theory and practice, requires strong assumptions on the form of one or both latent processes.

Contributions and roadmap. Our research makes three core contributions. First, we develop an asymptotic covariance analysis for the Pogit model and use it to qualitatively describe the limitations of the standard Pogit approach, which we illustrate using numerical examples. Second, we introduce new constraints and priors to robustify the estimation process, and we improve the ability to deconvolve under-reporting from the true count process using rigorous sandwich estimation to evaluate the uncertainty of model estimates. Finally, we develop an open source implementation of the algorithm and illustrate its successful application on a large-scale dataset that has a gold standard, validating our results.

Chapter organization. Section 4.2 describes existing models for under-reported counts. We develop an asymptotic analysis and numerical examples highlighting challenges of the deconvolution problem in Section 4.3. In Section 4.4, we present new methods to overcome these challenges using priors and constraints as well as novel approaches for robust uncertainty quantification. We develop and analyze case studies using real data in Section 4.5 and present a brief concluding discussion in Section 4.6.

4.2 Current models for under-reported counts

The standard approach to modeling under-reported counts assumes a two-step data generating process. Let Y_i represent the reported number of counts for observations $i = 1, 2, \ldots, n$. In the first step, the unobserved true number of events Y_i^* is drawn according to a Poisson distribution:

$$Y_i^* \sim \text{Poi}(\lambda_i) \tag{4.1}$$

In the second step, these events are filtered through a reporting process, where event i has probability p_i of being reported. The reported counts are modeled as a binomial random variable:

$$Y_i \sim \text{Binom}(Y_i^*, p_i) \tag{4.2}$$

This is equivalent to drawing the reported counts Y_i from the Poisson distribution

$$Y_i \sim \text{Poi}(\lambda_i p_i). \tag{4.3}$$

Estimating the reported process mean $\mu_i = \lambda_i p_i$ and its underlying process can be simply done using Poisson regression and the corresponding generalized linear model. However, in the under-reported counts setting, we need accurate estimates of both the true rate of events λ and the reporting rate p. Previous work proposed several models to separate (deconvolve) p and λ using observations of their product μ.

4.2.1 Models for the true rate λ and reporting rate p

The only way to deconvolve μ_i into factors $\lambda_i p_i$ is to incorporate additional assumptions about each component; otherwise, $\lambda_i = \mu_i$ and $p_i = 1$ is always a valid solution. Previous work on models for under-reported counts used distributional assumptions on parameters depending on whether the observations are associated with covariates, as described below.

Modeling without covariates

In the absence of covariates, a popular approach is to adopt a hierarchical model, where λ and p are latent variables drawn from underlying prior distributions. When a Gamma prior is placed on λ and a Beta prior on p, the resulting model is called the Beta-Binomial/Negative Binomial distribution, described and analyzed in Schmittlein et al., 1985. The specific choice of priors makes it tractable to compute posterior estimates of the individual λ_i and p_i, as derived by Fader and Hardie, 2000 for Empirical Bayes estimation of individual λ_i and p_i.

Modeling with covariates

Given covariates $x_{i,\lambda}$ that predict the true rates and covariates $x_{i,p}$ that predict reporting rates, we can model λ_i and p_i as functions of these covariates. The most popular model is the Poisson-Logistic regression, or *Pogit* model, proposed by Winkelmann and Zimmermann, 1993. In this model, the first step of the under-reported counts process (4.1) is modeled according to standard Poisson regression with coefficients θ_λ:

$$Y_i^* \sim \text{Poi}\left(\exp\left(x_{i,\lambda}^\top \theta_\lambda\right)\right). \tag{4.4}$$

The second step (4.2) is modeled according to logistic regression with coefficients θ_p:

$$Y_i \sim \text{Binom}\left(Y_i^*, \frac{\exp(x_{i,p}^\top \theta_p)}{1+\exp(x_{i,p}^\top \theta_p)}\right) \tag{4.5}$$

$$=: \text{Binom}\left(Y_i^*, \text{expit}\left(x_{i,p}^\top \theta_p\right)\right). \tag{4.6}$$

The generating distribution for Y_i can now be written as the Poisson distribution

$$Y_i \sim \text{Poi}\left(\exp\left(x_{i,\lambda}^\top \theta_\lambda\right) \text{expit}\left(x_{i,p}^\top \theta_p\right)\right). \tag{4.7}$$

The Pogit model has been used to estimate worker absenteeism in econometrics (Winkelmann, 1996), tuberculosis incidence in epidemiology (Stoner et al., 2019), and traffic accidents in highway engineering (Wood et al., 2016).

4.2.2 Parameter Estimation

Estimating the parameters of the Pogit model remains a deconvolution problem of λ and μ processes: a high observed count may be due to either a high underlying rate or a high reporting rate. Previous work addressed identifiability conditions for the Pogit model as well as several ways to include side information to improve the parameter estimates.

Conditions for parameter identifiability

A complete treatment of the difficulties in separating p and λ under the maximum likelihood framework is given by Papadopoulos and Silva (2012), who observed two distinct Pogit model parameterizations that lead to the same conditional law $\mathbb{P}(Y|x)$:

$$\mu_i := \exp(x_{i,\lambda}^\top \theta_\lambda) \frac{\exp(x_{i,p}^\top \theta_p)}{1+\exp(x_{i,p}^\top \theta_p)} = \exp(x_{i,\lambda}^\top \theta_\lambda + x_{i,p}^\top \theta_p) \frac{\exp(-x_{i,p}^\top \theta_p)}{1+\exp(-x_{i,p}^\top \theta_p)} =: \mu_i^a. \tag{4.8}$$

The authors show that identifiability can be regained either by knowing the sign of some nonzero element of θ_p a priori or by restricting some covariates to $x_{i,p}$, excluding them from $x_{i,\lambda}$. In earlier work (Papadopoulos and Santos Silva, 2008), the authors discussed problems that could arise if the restricted covariate is nearly colinear with the remaining covariates; in this case, they illustrated the resulting near-unidentifiability using employment data from the German Socio-Economic Panel (Wagner et al., 1993), which was previously analyzed in Winkelmann, 2008.

The identifiability problem (4.8) shows how the inherent ambiguity in deconvolving a product into individual terms directly translates to ambiguity in the Pogit model. In this work, we reduce this ambiguity by incorporating additional information in the parameter estimation, i.e., using constraints and regularization, to better resolve p and λ (and their associated models) in the maximum likelihood framework.

Variance reduction methods

Correlation among covariates for λ and p increases the risk of unidentifiability in the model and can manifest as high variance, as noted by Papadopoulos and Silva (2012). Another source of variance in the Pogit model is model misspecification, which can occur if there is overdispersion in the Poisson process. Several techniques have been developed to address these sources of variance.

One technique adds constraints or priors to the model, incorporating side information to reduce variance. This approach, used by Stoner et al. (2019), applies the Pogit model to the problem of estimating tuberculosis incidence in regions of Brazil using a Bayesian formulation. Here, the side information is a prior on the aggregate rate of tuberculosis reported across all regions, elicited from WHO estimates. The authors emphasize the strong dependence of the fitted model on this prior. Another type of side information useful for reducing variance is the presence of fully reported observations for which the reporting rate is one. This type of regularization induces the function p to pass through certain points and is used in the analyses of Stamey et al., 2006 and Dvorzak and Wagner, 2016.

A second technique addresses model misspecification by increasing the Pogit model's flexibility. Overdispersion of the true observations could be addressed by replacing the Poisson model with a negative binomial, although to our knowledge this has not been done previously. In the Bayesian setting, Stoner et al. (2019) address overdispersion by including additional Gaussian noise in the relationship between λ and p and their covariates.

4.3 *Characterizing the difficulties of p, λ deconvolution*

With the exception of Papadopoulos and Silva (2012), no work has analyzed the shortcomings of the Pogit model from a theoretical perspective. In Section 4.3.1, we derive an asymptotic lower bound for the variance of the maximum likelihood estimate of the Pogit parameters under a simplified setting, where p and λ each depend on a single covariate. This analysis reveals a fundamental difficulty: the variance of θ_p grows with θ_p^2, making it difficult to identify p, and hence λ, in a setting with moderate or large θ_p. We test this intuition using numerical simulations in Section 4.3.2, where we present a simple setting that nonetheless makes it impossible to infer even the sign of θ_p for any value of θ_p.

4.3.1 *Theoretical analysis in the two-covariate setting*

To characterize the behavior of the estimated Pogit model, we analyze a simple version of it. In our setting, p and λ are each determined by a single covariate

$$p_i = \frac{\exp(x_{p,i}\theta_p)}{1+\exp(x_{p,i}\theta_p)} \tag{4.9}$$

$$\lambda_i = \exp(x_{\lambda,i}\theta_\lambda) \tag{4.10}$$

so that

$$Y_i \sim \text{Poi}\left(\exp(x_{\lambda,i}\theta_\lambda)\frac{\exp(x_{p,i}\theta_p)}{1+\exp(x_{p,i}\theta_p)}\right). \tag{4.11}$$

We are interested in the performance of the estimators of parameters $\boldsymbol{\theta} = [\theta_p, \theta_\lambda]$, measured by the mean squared error

$$\text{MSE}(\hat{\theta}) := \mathbb{E}_{Y,X}\left[(\hat{\theta}-\theta)^2\right]. \tag{4.12}$$

For $i = 1, 2, \ldots, n$, let covariates $x_{p,i}$ and $x_{\lambda,i}$ be drawn independently according to

$$\begin{aligned} x_{\lambda,i} &\sim \mathcal{N}(\mu_\lambda, \sigma_\lambda^2) \\ x_{p,i} &\sim \mathcal{N}(0, \sigma_p^2) \end{aligned} \tag{4.13}$$

To analyze the behavior of the maximum likelihood estimates of θ_p and θ_λ, we assume as a further technical condition that $\theta_\lambda, \theta_p \in [C_l, C_u]$ for some constants $C_l, C_u \in \mathbb{R}$. These assumptions help us prove regularity conditions about the maximum likelihood estimator, but in practice they can be chosen as sufficiently large to be inactive at the solution.

First, we provide preliminary results about the Pogit model. We did not find these results in the literature and include them here for completeness, with proofs in Appendix C.1.

Recall that the Fisher information matrix is defined by

$$\mathcal{I}(\theta) = \mathbb{E}\left[(\nabla_\theta \log f(X,Y;\theta))(\nabla_\theta \log f(X,Y;\theta))^\top \middle| \theta\right], \tag{4.14}$$

where f is the probability density function of the data X and Y given parameters θ. In the supporting information, we show in Lemma C.1.1 that model (4.13) satisfies regularity conditions required for the MLE to be asymptotically normally distributed, and for the Cramér-Rao lower bound to hold. The inverse of the Fisher Information Matrix provides a lower bound on the variance of any unbiased estimator via the Cramér Rao lower bound; a lower bound on the Fisher Information Matrix therefore provides insight into the fundamental difficulty of the estimation task. Our main result is given below.

Theorem 4.3.1. *Let* $\{x_{\lambda,i}, x_{p,i}, Y_i\}_{i=1}^n$ *be drawn as in* (4.13)*. Let* $\hat{\boldsymbol{\theta}}$ *be any unbiased estimator of* $\boldsymbol{\theta} = [\theta_\lambda, \theta_p]$*. Then, the covariance of* $\hat{\boldsymbol{\theta}}$ *is lower bounded by*

$$Cov(\hat{\boldsymbol{\theta}}) \succeq \frac{1}{n\mathbb{E}[\lambda]} \begin{bmatrix} \frac{1}{\mathbb{E}[p]\left(\left(\mu_\lambda+\sigma_\lambda^2\theta_\lambda\right)^2+\sigma_\lambda^2\right)} & 0 \\ 0 & 2\theta_p^2 \end{bmatrix}. \tag{4.15}$$

To gain a more intuitive understanding of the result in Theorem 4.3.1, we take $\sigma_\lambda^2 = \sigma_p^2 = 1$ and $\mu_\lambda = 0$, in which case the asymptotic variance for $\hat{\theta_\lambda}_{MLE}$ and $\hat{\theta_p}_{MLE}$ can be lower-bounded by

$$\text{Var}(\hat{\theta_\lambda}_{MLE}) \geq \frac{2}{n\mathbb{E}[\lambda]\left(\theta_\lambda^2+1\right)} \tag{4.16}$$

$$\text{Var}(\hat{\theta_p}_{MLE}) \geq \frac{2}{n\mathbb{E}[\lambda]}\theta_p^2 \tag{4.17}$$

The standard deviation in estimating θ_p grows at least proportionally to the value of θ_p itself; for a fixed value of n, there will be at least a constant probability of incorrectly estimating the sign of θ_p, which is equivalent to incorrectly determining whether p is an increasing or decreasing function of x_p. In applications where estimating p is important, additional information and constraints *must* be used to decrease the variance of $\hat{\theta_p}_{MLE}$.

4.3.2 Simulation Study

We perform a simulation study to illustrate the high variance of p and λ in this simple setting. For our simulations, we generate $n = 50$ covariates $x_\lambda \sim \mathcal{N}(0,1)$ and $x_p \sim \mathcal{N}(0,1)$.

For the first simulation, shown in Figure 4.1 (left), we fix $\theta_\lambda = 1$ and vary θ_p from -5 to 5, then we report the 95% confidence interval based on the empirical standard deviation of $\hat{\theta}_p$ and $\hat{\theta}_\lambda$ over 5000 replicates. We compare this to our theoretical lower bound on the confidence interval for each parameter. Our experiments confirm the theoretical bound in this setting: the standard deviation in $\hat{\theta}_\lambda$ is independent of θ_p, while the standard deviation of $\hat{\theta}_p$ grows linearly with the parameter θ_p, highlighting the difficulty of estimating p in this setting. We also see that $\theta_p = 0$ is always within the 95% confidence interval, indicating that no matter how large θ_p truly is, we cannot determine whether p increases or decreases with x in this setting.

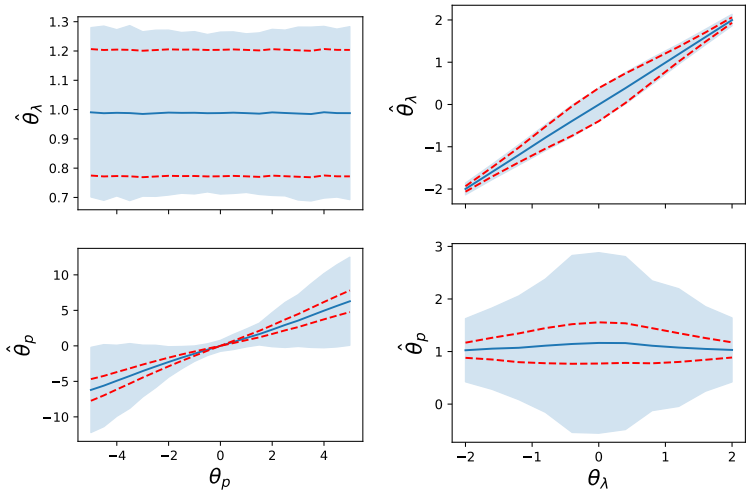


Figure 4.1: Simulation studies illustrating the dependence of $\text{Var}(\hat{\theta}_{p,MLE})$ and $\text{Var}(\hat{\theta}_{\lambda,MLE})$ on θ_p and θ_λ for $n = 50$ data points. The left panels vary θ_p with fixed $\theta_\lambda = 1$, while the right panels vary θ_λ with fixed $\theta_p = 1$. Solid blue lines show the average parameter estimate over 5000 trials; the blue shaded region is the 95% confidence interval based on the empirical standard deviation. The red dashed lines show our theoretical lower bound on the 95% confidence interval.

We perform a second simulation with $\theta_p = 1$ fixed while θ_λ varies from -2 to 2, shown in Figure 4.1 (right). The results confirm our interpretation of the lower bound; the standard deviation in both $\hat{\theta}_\lambda$ and $\hat{\theta}_p$ decreases as the magnitude of θ_λ increases.

These simulations illustrate that the estimated parameters can suffer from high variance; in particular, we can never reject the hypothesis that $\theta_p = 0$ for any value of θ_p. In practical settings, we need additional information to reduce the variance in the model and successfully deconvolve p and λ. We next discuss a framework for incorporating priors and constraints as a powerful way to bring expert knowledge to bear on specific problems.

4.4 Incorporating prior knowledge into model building

We now discuss how we use prior knowledge to deconvolve the under-reporting and data-generating mechanisms while maintaining model identifiability (Section 4.4.1) and how we apply the sandwich estimation procedure to quantify the resulting model's uncertainty (Section 4.4.2).

4.4.1 Covariate Specifications

To deconvolve the under-reporting mechanism (parametrized by p) from the data-generating mechanism (parametrized by λ), we model both as functions of covariates. These functions consist of a linear predictor and a link function that transforms the linear predictor to the desired space, as

in generalized linear models. For example, to model the parameter λ using a vector of covariates $\boldsymbol{x}$, a vector of regression coefficients $\boldsymbol{\theta}$, and link function g:

$$\lambda := g^{-1}(\boldsymbol{x}^\top \boldsymbol{\theta}).$$

Linear predictors. For $\boldsymbol{x}^\top \boldsymbol{\theta}$, we can choose simple covariate specifications, like including a continuous predictor as a single linear term. Alternatively, nonparametric regression techniques (such as basis splines (De Boor, 1978)) let us flexibly parametrize the relationship between a covariate and λ or p. These spline specifications, which include the degree of the spline and the number and location of knots, are embedded in the linear predictor $\boldsymbol{x}^\top \boldsymbol{\theta}$.

To encode knowledge about the shape of the relationship among covariates, we can use linear inequality constraints on the regression coefficients $\boldsymbol{\theta}$. For example, such constraints could force the regression coefficient to be positive, which would enforce an increasing relationship between the true rate of reporting and a given covariate. Further, very general linear constraints are particularly useful for working with basis splines. The second derivative of a basis spline can be represented as a linear function of its basis elements, so linear constraints of the regression coefficients let us constrain the second derivative to be positive or negative.

Finally, rather than constraining the regression coefficients $\boldsymbol{\theta}$ of the linear predictor, we can include a quadratic regularizer, commonly known as a Gaussian prior or 'ridge' regression penalty (Hoerl and Kennard, 1970). Trading off bias for variance in the parameter estimation, we can avoid over-fitting the data, and we can incorporate prior beliefs in a quantitative way.

Link functions. In the Pogit model, we use different link functions g for λ and p. Specifically, we use the logit function for p and the log function for λ. W can also use other functional forms,as needed. For example, if we understand from prior knowledge that the under-reporting rate is between a and b, the inverse link function

$$l_p(x) = a + \frac{b-a}{1+\exp(-x)}$$

for parameter p captures this information with no need for more complex constraints.

Example. To show the impact of the innovations on the pogit model, we use a simple synthetic example, where p and λ are taken to be simple nonlinear functions

$$\lambda = 15 + \exp\left(\cos(2\pi x_0)\right), \quad p = \text{expit}\left(\sin(2\pi x_1)\right),$$

with x_0 and x_1 simulated as independent uniform random variables. The Pogit model is fitted using only reported data. Spline specifications for p and λ are used to capture the nonlinear relationships. Figure 4.2 shows the results for predicted p, λ, and $\mu = p\lambda$ across 100 realizations of the experiment. Its first column presents results for the unconstrained spline Pogit approach; though the μ fit is correct (third row), resolving p and λ is far more difficult. In each column thereafter, we show

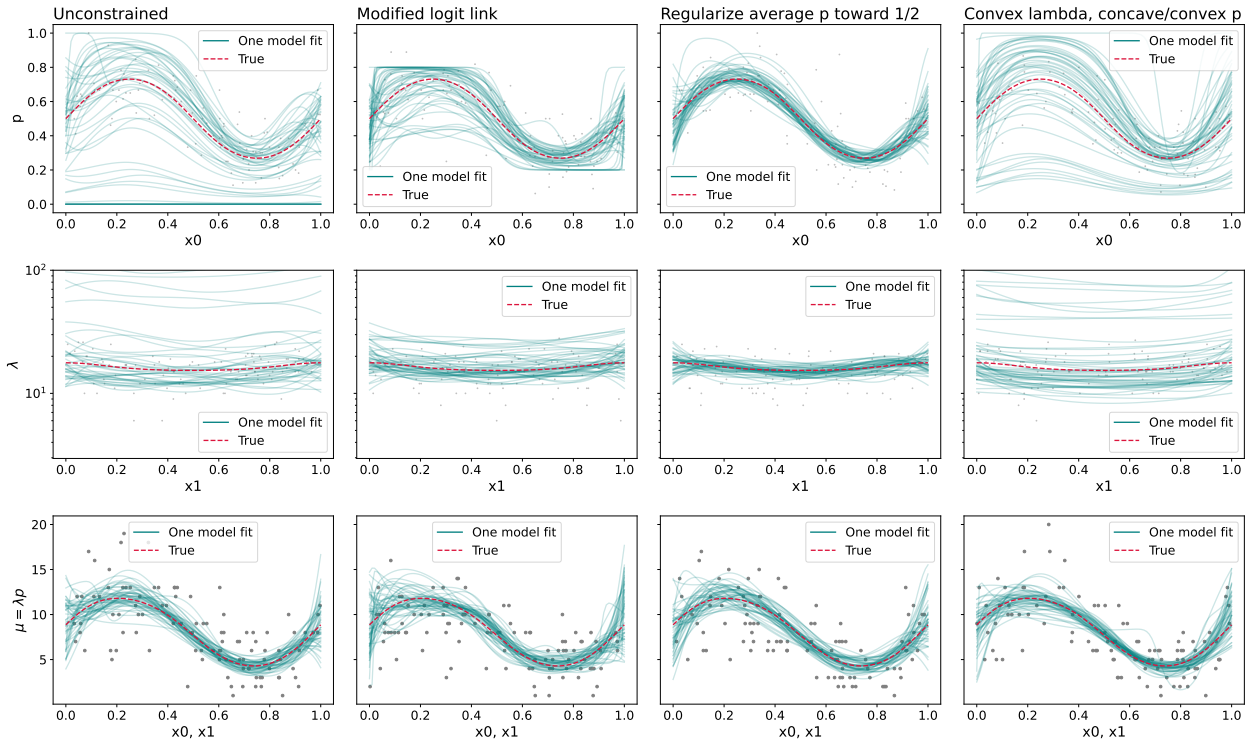


Figure 4.2: Improvements in fits vs. baseline pogit model (first column) of: modified logit link (second column), prior on p (third column), and convexity constraints (fourth column). All methods easily fit μ (third row) since it is directly informed by the observations. All three innovations improve fit for p (first row) and λ (second row), with the prior on p showing the largest impact in this example.

the impact of the three techniques described above. The table's second column shows results for the modified link function that bounds p between 0.2 and 0.8 through its representation. The third column shows results for using quadratic regularization pulling p to 0.5. Finally, the fourth column presents imposing convexity constraints on p (as a function of x_1) and λ (as a function of x_0). All three techniques improve resolution of p and λ, with quadratic regularization helping the most: it provides specific (strong) information about the value of p rather than general (weaker) information about the bounds on p or the nature of the relationship between p and x_1 or λ and x_0. All approaches are comparable in their recovery of μ, which underscores the fact that the μ fit alone cannot differentiate successful resolution of p and λ (e.g., as in column 3) and failure to resolve these parameters (as in the unconstrained results of column 1).

4.4.2 *Uncertainty quantification and model diagnostics*

We use a robust approach to uncertainty quantification that allows estimation of both the uncertainty for p, λ, μ as well as covariate multipliers describing the relationships between these parameters and covariates. Sandwich estimation (Kauermann and Carroll, 2001; Wakefield, 2013) is robust to model misspecification, which proves to be particularly important for the Pogit model. Specifically, our variance-covariance matrix is computed as

$$V = A^{-1}BA^{-1},$$

where A is the Hessian of the log likelihood $\mathbb{E}_\theta[\nabla^2_\theta \ell(x, y|\theta)]$, and B is the Gauss-Newton Hessian approximation $\mathbb{E}_\theta[\nabla_\theta \ell(x, y|\theta) \nabla_\theta \ell(x, y|\theta)^\top]$, both computed at the maximum likelihood estimate by their empirical approximations of the expectations. In practice, this approach reports wider uncertainty intervals in the presence of model mis-specification, helping modelers to detect difficult cases.

4.5 *Case studies: validation on injury datasets*

We present two case studies exploring the performance of the Pogit model on health-related datasets. Our two case studies, on interpersonal violence and diabetes, illustrate the use of prior knowledge in the form of covariates, constraints and regularization to address the challenges of identifiability and high variance (described in Section 4.3). In the interpersonal violence study, we estimate the rate of injuries warranting medical care using data from injuries warranting only *inpatient* medical care, allowing us to apply the "under-reporting" framework. For the diabetes study, we estimate the overall rate of medical encounters, again using only inpatient data. For each case, we validate our predictions from the Pogit model by comparing them to the total of inpatient *and* outpatient data.

4.5.1 *Case Study: Interpersonal Violence*

In the International Classification for Diseases 9 (ICD-9) and 10 (ICD-10) codes, injuries are classified by their cause (e.g., interpersonal violence) and/or their nature (e.g., traumatic brain injury)Vos et al. (2020). In addition, they are reported separately based on outpatient or inpatient status. For this case study, we consider all injuries resulting from interpersonal violence and separate them by treatment inside or outside the hospital setting.

For the validation setting, λ is the true rate of all inpatient and outpatient injuries combined, p is the proportion of injuries that are seen in the inpatient setting, and μ is the observed rate of injuries in the inpatient category. Our goal is to recover the total rate of injuries due to interpersonal violence *from only inpatient information.* We use inpatient and outpatient interpersonal violence injury data aggregated at the national level for the US by The Global Burden of Disease study (National Center for Health Statistics and Prevention, 2018). Only two covariates, age and sex, are available in this dataset.

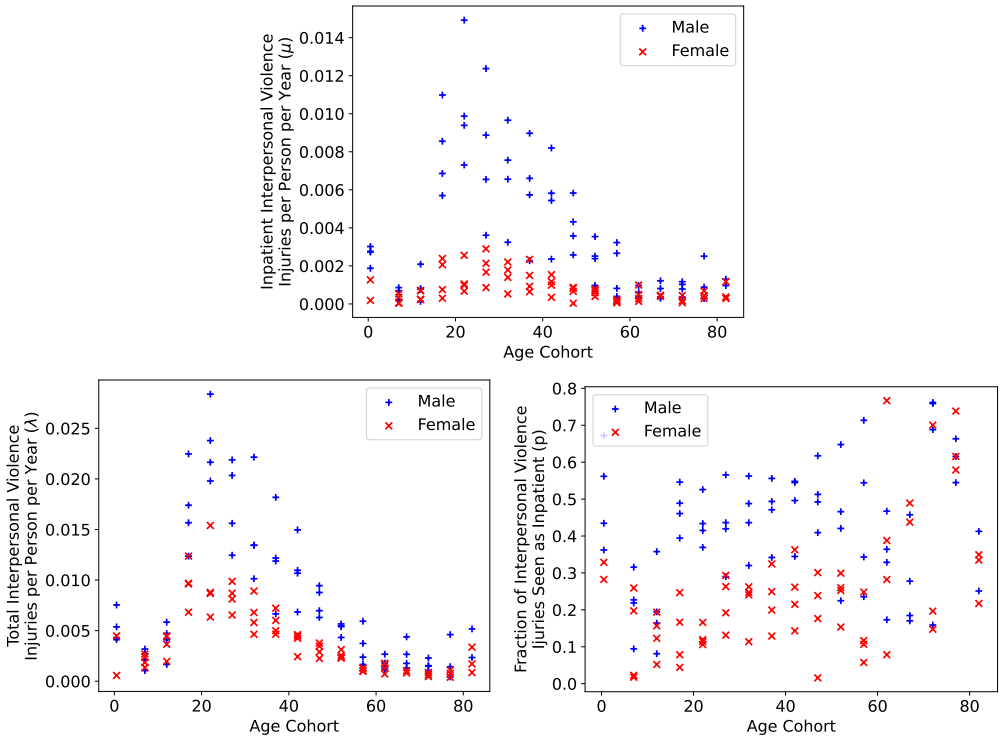


Figure 4.3: Top: Observed (inpatient) interpersonal violence injuries in the United States, plotted by age/sex, over four five-year periods between 1993 and 2012. Bottom left: true plot of λ, the combined rate of inpatient and outpatient injuries. Bottom right: true plot of p, the rate at which injuries are treated as inpatient. The Pogit model sees only inpatient data.

Interpersonal violence illustrates an interesting case for our models: the covariates controlling the reporting rate are a strict subset of the covariates controlling the true rate, leading to a loss of identifiability that can be recovered only by using constraints. In particular, as explained below, we model the rate of injury as a function of age and sex and the probability of inpatient care using only sex.

Figure 4.3 shows the observed (inpatient) injury rate per person per year, split by age and sex cohorts, for five-year periods between 1993 and 2012. There are clear age and sex effects in the data, with young adult males having the highest rates of injuries requiring inpatient care.

Domain knowledge is a critical component of the modeling process, and identifiability of the Pogit model depends on proper modeling choices for p and λ. In both of our case studies, we use plots of true p and λ, shown in Figure 4.3, to make reasonable choices for their functional forms. Based on these plots, we model λ as a spline in age and sex and p as a function of sex alone. This information would not be available to modelers in the real under-reporting setting, who would need domain knowledge to determine the functional forms of p and λ.

Modeling λ as a function of age and sex.

We model the true injury rate, λ, as a function of age a_i, sex s_i (coded 0 for males and 1 for females) and a fitted intercept:

$$\lambda_i = \exp(\beta_{\lambda,0} + \beta_{\lambda,1} s_i + f_\lambda(a_i)). \tag{4.18}$$

Age enters the model as a cubic spline f_λ with a knot at age 15. The placement of knots can be guided by domain knowledge, e.g., about change points in interpersonal violence based on age.

Modeling p as a function of sex alone.

As Figure 4.3 shows, the fraction of interpersonal violence requiring inpatient care (the fraction "reported") is primarily a function of sex, with a higher rate for males than females. We model this as

$$p_i = \frac{\exp(\beta_{p,0} + \beta_{p,1} s_i)}{1 + \exp(\beta_{p,0} + \beta_{p,1} s_i)}. \tag{4.19}$$

Constraints and regularization.

The covariates modeling p are a subset of those used to model λ. As discussed by Papadopoulos and Santos Silva (2008) and in Section 4.2.2, this overlap in covariates renders the model unidentifiable. To recover identifiability, we add the constraint that males seek inpatient care at higher rates than females. This constrains $\beta_{p,1} > 0$ in Eqn (4.19) and results in an identifiable model.

Even with this constraint, substantial variance remains in the model predictions for p and λ. Our approach to further reduce this variance is to add regularization for p, pushing p toward 0.5 using an ℓ_2-norm penalty on the magnitude of $\beta_{p,0} + \beta_{p,1} s_i$.

Results.

Even with regularization and constraints, the variance of model predictions remains high. Figure 4.4 shows the fitted model and its components $\hat{p}$ and $\hat{\lambda}$. We see that (1) the fitted $\hat{p}$ is higher than the true reporting rate for both sexes, and (2) the variance is so high that the confidence intervals span almost the entire range $[0, 1]$, signalling that the problem is difficult. Nonetheless, we can still recover a reasonable estimate for $\hat{\lambda}$, which is typically the more important quantity from a global health perspective.

We obtain a quantitative comparison of our $\hat{\lambda}$ estimate vs. baseline estimates using the Akaike Information Criteria (AIC). The AIC is given by $2k - 2\log\hat{\ell}$, where ℓ is the likelihood of the model on the fully reported data Y_i^* (see Equation (4.2)) and k is the number of parameters. The first model we compare to is the *oracle model* of λ. The oracle observes the combined inpatient and outpatient data (Y_i^*, in the notation of Equation (4.1)) and fits the Poisson model (4.18) to that data. This represents the maximum likelihood fit to λ within the model class of Equation (4.18). Our second comparison is to the naive baseline of *ignoring under-reporting*. For this baseline, the

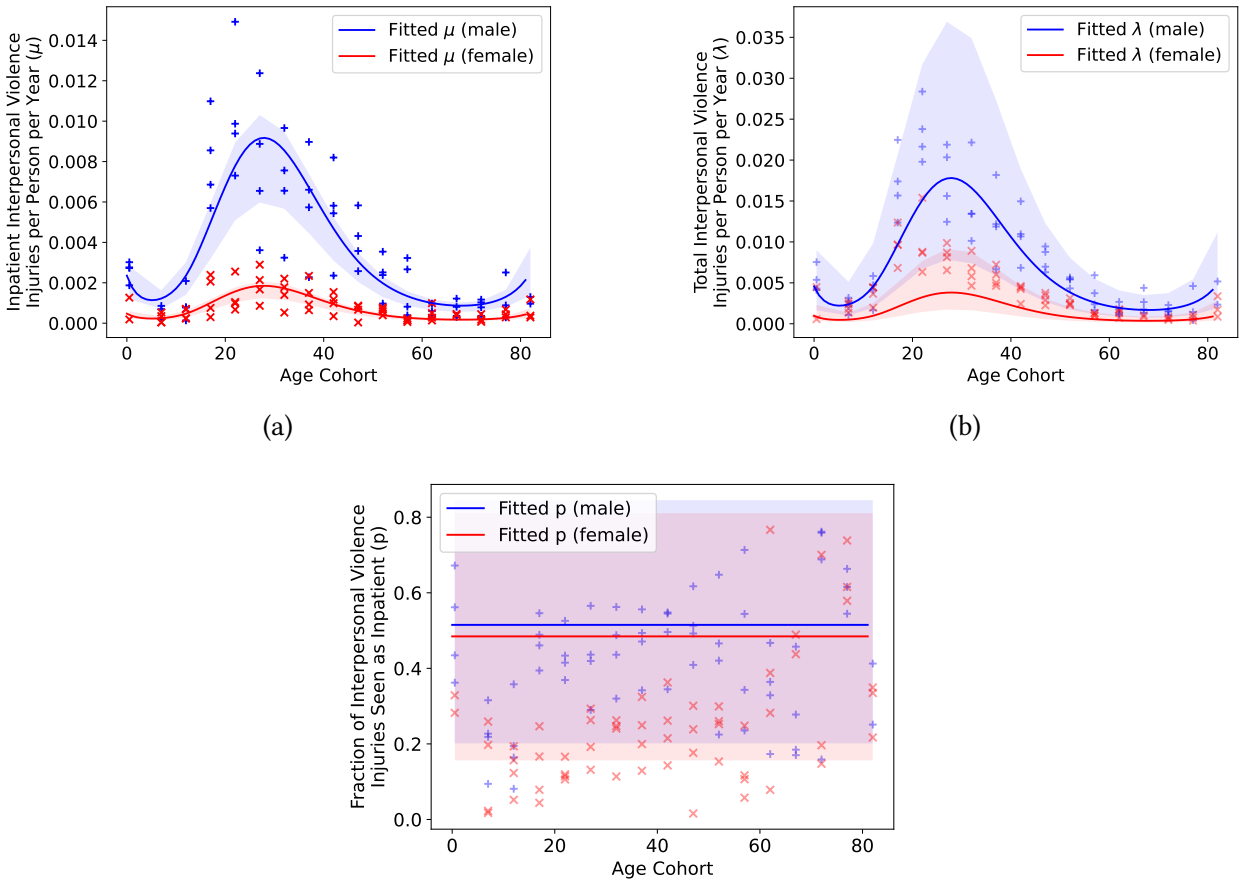


Figure 4.4: Regularized Pogit fits for the interpersonal violence injuries model using only inpatient data. The estimated total rate λ and fraction inpatient p are plotted for each age/sex cohort against validation data not available to the model. Shaded intervals are 90% confidence intervals computed using sandwich estimation, as described in Section 4.4.2.

Table 4.1: AIC of the true rate of events λ over all data points Y_i^*, reported for the Poisson model (4.18) for injuries and (4.20) for diabetes, under three different choices of training data: the "oracle" Poisson fit of the injury rate to the Poisson model if we could observe the number of total events Y_i^*; the Poisson portion of the Pogit model fit to observed events Y_i; and the naive baseline that ignores under-reporting by fitting the Poisson model to the observed events Y_i.

	Oracle Fit	Pogit Fit	Ignoring Under-reporting
Injuries	18000	62000	156000
Diabetes	450	749	37400

Poisson model (4.18) is fit to the inpatient data only (Y_i in the notation of Equation (4.2)). This baseline, which represents the model that is unaware of the under-reporting problem, will have a low likelihood on the true Y_i^* when the observations are severely under-reported.

Table 4.1 shows the AIC values for the three models we consider. Since all three models are of the same parametric form, the $2k$ term acts as a constant offset. We see that the Pogit fit is significantly better compared to ignoring under-reporting and significantly worse than the oracle fit ($p < 10^{-10}$ with a likelihood ratio test in both cases).

4.5.2 Case Study: Diabetes Care

In the second case study, we apply the Pogit model to estimate the rate of diabetes care (inpatient *and* outpatient visits) having observed *only inpatient* visits. We use MarketScan healthcare claims data that is processed for use in the Global Burden of Disease Study (Analytics, 2010; Vos et al., 2020; National Center for Health Statistics and Prevention, 2010). The data is aggregated at the age and sex level for each US state for the year 2019; the state-level aggregation lets us use a richer set of covariates to model p and λ. Figure 4.5 shows the rate of inpatient diabetes cases per person per year across all fifty states, as a function of age, sex, and population average fasting plasma glucose (FPG) for each state/age/sex cohort. Both age and population average FPG correlate positively with diabetes inpatient admissions. We parametrize models for p and λ based on the p and λ plots shown in Figure 4.5.

Modeling λ as a function of FPG and sex.

We model the true rate of diabetes care, λ, as a function of sex s_i and the state average FPG g_i with a fitted intercept:

$$\lambda_i = \exp(\beta_{\lambda,0} + \beta_{\lambda,1} s_i + f_\lambda(g_i)). \tag{4.20}$$

FPG enters the model as a quadratic spline f_λ.

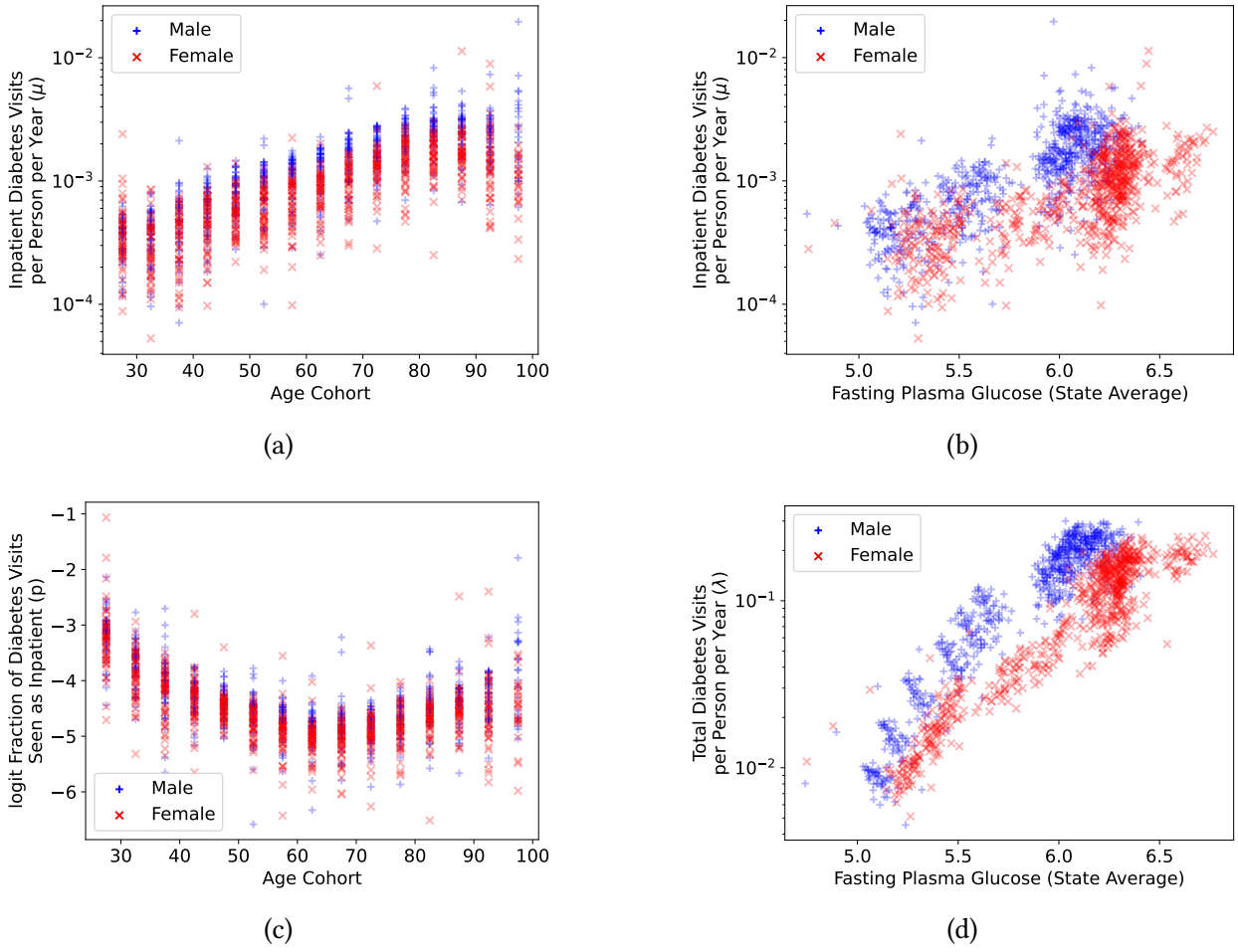


Figure 4.5: Inpatient diabetes care for age/sex (a) and cohort-average FPG (b) in 2019. True fraction of inpatient visits as a function of age and sex for 2019 (c). Total diabetes care as a function of population average FPG and sex for 2019 (d).

Modeling p as a function of age.

The observed rate of inpatient diabetes care is driven by the true rate of diabetes care and by the fraction of care that is treated in the inpatient vs. outpatient setting. Based on Figure 4.5, we model p as a quadratic spline in age:

$$p_i = \frac{\exp(\beta_{p,0} + f_p(a_i))}{1 + \exp(\beta_{p,0} + f_p(a_i))}. \tag{4.21}$$

We apply several constraints and regularizers on p to reduce the variance of our estimate. First, we enforce that p decreases from age 25 to age 60. Second, we apply a quadratic regularization of the average fitted p toward its true average, 0.01. Including this side information improves the model fits for both p and λ.

Results

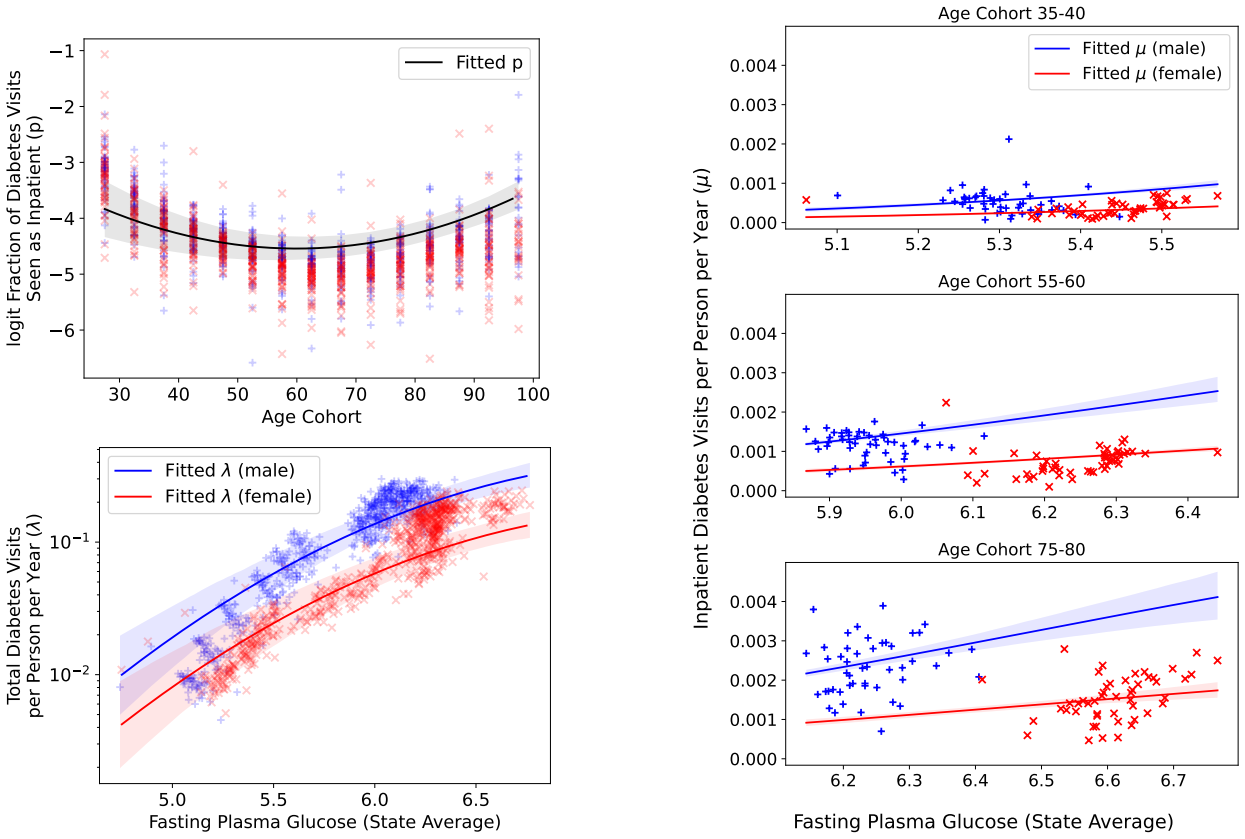
We fit the model described in (4.20) and (4.21) to the diabetes inpatient care data. Figure 4.6 shows the results. Both fitted $\hat{p}$ and $\hat{\lambda}$ capture the important properties of their respective processes: p is convex in age, while λ is concave in FPG and shows a higher rate for males than females.

Following the methodology of Section 4.5.1, we quantitatively evaluated the fitted model of the combined rate of care, λ, using the AIC. Table 4.1 shows the AIC of the Pogit model, the oracle with access to both inpatient and outpatient data, and the naive baseline that ignores under-reporting. The Pogit fit is significantly better than the naive baseline, but not as good as the model that observes both inpatient and outpatient data ($p < 10^{-10}$ with a likelihood ratio test in both cases). The Pogit model outperforms the naive baseline by a wider margin on the diabetes data than on the interpersonal violence study because the reporting rate is lower for diabetes, so the penalty for ignoring under-reporting is higher.

4.6 Discussion

In this chapter, we presented theoretical challenges in modeling under-reported count data using the Pogit model. We showed how priors and constraints can help resolve these issues and used real-world data from the Global Burden of Disease study to validate our approach. We found that the proposed formulation enables successful estimation, provided sufficient prior information can be specified. Examples of such information include aggregate measures (such as national reporting rate), shape of the relationships, and prior values for specific datapoints and covariate values. The tools used to create the results and test the methods are available in a publicly accessible repository.

The approach and analysis in this chapter focused on the Pogit model. Future analysis and extensions can be made by considering other count models that better account for over-dispersion. A potential challenge for extensions is that additional flexibility may exacerbate the difficulty of the deconvolution problem.



(a) (b)

Figure 4.6: Regularized Pogit fits for the diabetes care model using only inpatient data. The estimated total rate of care λ and fraction of inpatient care p are plotted for each age/sex cohort against validation data that was not available to the model. The fitted inpatient care rate μ is a function of age, sex and state average FPG, so we provide three plots for different age cohorts. Shaded intervals are 90% confidence intervals computed using sandwich estimation, as described in Section 4.4.2.

Another interesting direction for future work is to use the approach developed here to aid in decision making about the kinds of data or information in which to invest. This idea is expanded in the next section.

4.6.1 Future Work: Prioritizing survey data collection using the principles of experimental design

We have seen that regularization is an important tool to effectively estimate the parameters θ_λ and θ_p of the Pogit model. However, not all types of regularization are equally informative. For example, a Bayesian study modeling spatiotemporal patterns in under-reporting of tuberculosis in Brazil Stoner et al. (2019) found that the results were highly sensitive to the prior on the average rate of reporting at a national level (estimated from World Health Organization country-level data), but less sensitive to other priors used in the model. In global health settings, priors/regularization can typically be elicited by either commissioning surveys or purchasing data from third parties. A natural question then becomes: which data is the most valuable, in the sense of providing the greatest uncertainty reduction in our model estimates?

To answer this question, we propose using tools from experimental design: namely, control of the Fisher Information Matrix. Consider adding ℓ_2 regularization to certain linear combinations of the parameters θ, where Z denotes the linear combination, r denotes the value toward which we will regularize, and σ controls the strength of the regularization (evocative of a Gaussian prior). As an example, if we had collected data on the population average reporting rate, then Z would encode the population average value of each covariate X_p, the target r would be the expit of the average reporting rate, and σ would encode the uncertainty in our estimate. Other constraints, such as information about the average reporting rate across a subset of the population, could also be used. If $\mathcal{L}$ is the likelihood function for the Pogit model, then the regularized objective function with m constraints can be written as

$$\arg\min_{\theta} -\mathcal{L}(\theta|X) + \sum_{i}^{m} \tfrac{1}{2}\sigma_i^{-2}(Z_i^T\theta - r_i)^2. \tag{4.22}$$

Constructing the optimal design (the optimal set of m constraints) involves controlling the shape of the Fisher information matrix, which in expectation is the inverse of the Hessian of the (regularized) log likelihood function in (4.22). Building off of our computation of the Hessian of the unregularized objective, we have

$$H(\theta) = \mathbb{E}[e^{X_\lambda\theta_\lambda}X_\lambda^2]\mathbb{E}\left[\tfrac{\exp(X_p\theta_p)}{1+\exp(X_p\theta_p)}\right] + \sum_{i}^{m}\sigma_i^{-2}Z_iZ_i^T. \tag{4.23}$$

Recall that, outside of the special case of linear models, optimal experimental design is complicated by the fact that the Fisher information matrix depends on the true, but unobserved, parameters θ. This is the case in the Pogit model as well. In such cases, we might choose to perform experimental designs over several rounds to refine our estimates of θ, or to start with a prior on θ; these ideas are expanded upon in section 2.1.4 of Temkit (2014).

Given an initial model estimate $\widehat{\theta}$, we observe that the Hessian (4.23) depends on the constraints only through the structure of the information gathered, Z, and the uncertainty in its estimate, σ, but not through the estimated value r. This is good news because Z and σ can be evaluated before data acquisition. For example, in the global health context, one may purchase a database of ground truth data aggregated by age and sex (encoded by Z) with a stated sample size (encoded by σ). It is important to note that the empirical Fisher Information matrix is not completely independent of the choice of constraints: since $\widehat{\theta}$ is the solution to a constrained optimization (4.22), the choice of constraints will indeed inform $\widehat{\theta}$, leading to the suboptimality of single-round design. However, our hope is that the choice of optimal constraints will be pursued as a refinement after an initial Pogit model has been fitted, when the estimate $\widehat{\theta}$ is accurate enough to provide a good approximation of the true Fisher information matrix.

When the constraints (Z, σ, r) can be collected as part of commissioned surveys, instead of fixed data sets for purchase, we can also ask for the optimal allocation of a limited sampling budget to various surveys. If $\sigma_i \propto \sqrt{B_i}$ scales with the sampling budget B_i assigned to each possible survey Z_i, we could define a cost-sensitive version of the problem. Instead of choosing a fixed number m of constraints to elicit, we could choose a fixed budget B for the number of people to survey.

Finally, we observe that similar ideas have been explored in the Bayesian literature under the name "Value of Information" (see, for example, chapter 13 of Parmigiani and Inoue (2009)). While this work is related to ours, it exists within the Bayesian framework, and does not connect to the ideas of Fisher Information. Our work also relates to ideas for experimental design in generalized linear models, particularly to the approach of iteratively performing optimal design based on the empirical Fisher information matrix (so-called "locally optimal design" - see section 2.1.4 of Temkit (2014) and references therein). We note that the Pogit model is not a generalized linear model, and that we believe the distinction between collecting more observations X and eliciting constraints Z brings a new perspective to this problem.

4.7 Software availability

The techniques described in this chapter have been implemented in the Python package `Regmod`, available on GitHub at `https://github.com/ihmeuw-msca/regmod`. Tutorials are available on GitHub at `https://github.com/ihmeuw-msca/underreporting`.

Chapter 5

Cluster Randomized Designs for One-Sided Bipartite Experiments

The conclusions of randomized controlled trials may be biased when the outcome of one unit depends on the treatment status of other units, a problem known as *interference*. In this chapter we study interference in the setting of one-sided bipartite experiments, a common model in marketplaces and two-sided platforms.

5.1 Introduction

Interference is a well-studied phenomenon in causal inference, whereby the treatment status of one unit can affect the outcome of another. Formally a violation of the Stable Unit Treatment Value Assumption (Rubin, 1980), interference has been studied in many settings, including agricultural studies (Neyman, 1923), clinical trials (Hudgens and Halloran, 2008), social networks (chap. 16 of Aronow et al., 2021; Aral and Walker, 2011; Gui et al., 2015; Eckles et al., 2016), and marketplaces (Blake and Coey, 2014; Rolnick et al., 2019; Li et al., 2022; Wager and Xu, 2021; Munro et al., 2021). Marketplaces exhibit unique forms of interference because they involve two types of units: *buyers* and *sellers*. Units of the same type do not interact directly, but rather their interactions are mediated through their interactions with units of the opposite type. For example, when buyers compete to buy limited goods, an increase in the price one buyer is willing to pay for a good will affect the market-clearing rate for that good, thus increasing the price for all other buyers.

Marketplace experiments can be conceptualized as running an experiment on a bipartite graph between buyers and sellers, with edges representing buyer-seller interactions. Other settings beyond marketplaces can also be formalized as bipartite graphs, such as content platforms (matching viewers to creators) or ride-sharing apps (matching riders to drivers) (Chamandy, 2016). We consider perhaps the most straightforward way of running an experiment on a bipartite graph: by treating and measuring a single side of the graph, a setting we call the *one-sided bipartite experiment framework*. We refer to the units on this side of the graph as *experimental units* since they receive treatment and their outcomes appear in our estimators. We refer to the units on the other side of the graph as *interference units*: while they do not explicitly receive treatment, they mediate the interactions between experimental units.

The presence of interference units distinguishes one-sided bipartite experiments from other settings of network interference because the interactions between experimental units are known in the latter case, but must be inferred from the interactions with interference units in the former.

In this chapter we show that a popular class of experimental designs for network interference, *cluster-randomized designs*, does not extend immediately to this setting. Instead, we propose a variant of the cluster-randomized design, with a clustering objective that accounts for the way in which interference units mediate the interference between experimental units. We motivate the use of the difference-in-means estimator in this setting and show that clustering according to our objective minimizes the bias of that estimator in a minimax sense over the class of linear potential outcomes models with bounded interference. We further illustrate the failure of existing clustering designs when applied to one-sided bipartite experiments, and conclude by illustrating the robustness of our design to a variety of potential outcomes models and bipartite graphs both theoretically and empirically.

5.1.1 Related work

Accurate estimation of the treatment effect requires some knowledge of the mechanism of interference (Basse and Airoldi, 2018); otherwise, having even a single treated (resp. control) unit in a graph of control (resp. treated) units could change the outcome of every unit arbitrarily. Existing work varies in the strength and nature of the assumptions on the interference model, but in general these assumptions take one of two types. The *exposure mapping* approach (Hudgens and Halloran, 2008; Manski, 2013), formalized by Aronow and Samii (2017), defines a notion of when a unit is "completely treated" or "completely controlled," then uses the inverse propensity score (IPS, also known as Horvitz-Thompson) estimator to construct an unbiased estimate of the average treatment effect. By contrast, an alternate approach is to propose a model for the effect of interference on the potential outcomes, and then rely on this model to estimate the average treatment effect using data from all units (even the ones experiencing a great deal of interference) (Gui et al., 2015; Eckles et al., 2017; Pouget-Abadie et al., 2019; Harshaw et al., 2022). Since the quality of the estimates depends on the accuracy of the modeling assumptions, several methods have been developed to estimate the magnitude and form of interference (Aral and Walker, 2011; Saveski et al., 2017; Aronow, 2012; Athey et al., 2018; Toulis and Kao, 2013; Sävje et al., 2021). Even in cases when the model is only approximate, so that including data from units experiencing interference may bias the estimator, Eckles et al. (2017) use realistic graph models to argue that the bias incurred is more than offset by the reduction in variance achieved by avoiding an IPS estimator. Our work assumes this same regime, in which the graphs are sufficiently dense that an IPS estimator will be too high-variance to be practical, so we accept an estimator with some bias. We assume a linear model of the potential outcome on the measured exposure, as in Gui et al. (2015) and Harshaw et al. (2022). To avoid strong dependence on the linear assumption, we use the difference-in-means estimator which does not rely on the model of interference, and we choose our experimental design to be minimax optimal over the class of linear potential outcomes models.

Given some model of interference and a choice of estimator, the next question is how to design an experiment (an assignment of units to treatment or control). The most popular experimental design in the case of network interference is the cluster-randomized design, studied by Ugander et al. (2013), Eckles et al. (2017), and Candogan et al. (2021) in the case of non-bipartite graphs. This

design first clusters units according to the provided graph, and then assigns each cluster to either treatment or control. In the specific case of bipartite graphs, other works propose modifications to unit-level randomization by choosing which side of the graph to randomize (Johari et al., 2022; Li et al., 2022; Bajari et al., 2021). Pouget-Abadie et al. (2019) and Harshaw et al. (2022) look directly at clustered designs on bipartite graphs, but they study a two-sided experimental framework in which one side of the graph is randomized while the other is measured. Perhaps closest to our work is that of Rolnick et al. (2019), which suggests a balanced partitioning of geographical regions using a clustering objective that is similar to ours. Their work considers a more restrictive form of the potential outcomes model, and uses clustering heuristics that are specific to the geographical setting. By contrast, our work considers the robustness of the design to a broader class of potential outcomes models and extends beyond geographical regions to general interference graphs.

5.2 Models and estimators

We now formalize the interference model using the potential outcomes framework (Imbens and Rubin, 2015). Let $\mathbf{Z} \in \{-1, 1\}^N$ be an assignment of each of the N experimental units to treatment ($Z = 1$) or control ($Z = -1$). The *potential outcome* of the i^{th} experimental unit is denoted Y_i, and in the most general setting could be a function $Y_i(\mathbf{Z})$ of the entire treatment assignment vector. We further assume the existence of a (known) bipartite graph between experimental units and interference units, with nonnegative weights $w_{is} \geq 0$ encoding the relationship between experimental unit i and interference unit s. Such a graph may be obtained from historical data on interactions between units i and s, or on similarity between i and s measured by geography or other features (Rolnick et al., 2019; Zigler and Papadogeorgou, 2021).

While there are many possible estimands of interest, our primary goal is to estimate the *average total treatment effect*

$$\tau = \frac{1}{N} \sum_{i \in [N]} Y_i(\mathbf{Z} = \mathbf{1}) - Y_i(\mathbf{Z} = -\mathbf{1})$$

sometimes referred to as the average treatment effect or total treatment effect.

Since we assign some units to treatment and others to control, it is impossible to observe any potential outcome under a fully treated ($\mathbf{Z} = \mathbf{1}$) or fully controlled ($\mathbf{Z} = -\mathbf{1}$) condition. If the underlying interference graph among units were composed of multiple connected components, so that only the assignments Z in a unit's connected component affected its potential outcome, then it would be possible to assign treatment at the level of the connected component and observe fully treated or control outcomes. However, in realistic marketplaces, such a perfect separation almost never occurs. As a result, we require some modeling assumptions on the potential outcomes to infer the behavior of unit i under the fully treated or controlled condition.

5.2.1 The potential outcome model

A popular approach to modeling potential outcomes with network interference is the *exposure mapping* paradigm (Hudgens and Halloran, 2008; Manski, 2013; Aronow and Samii, 2017). An

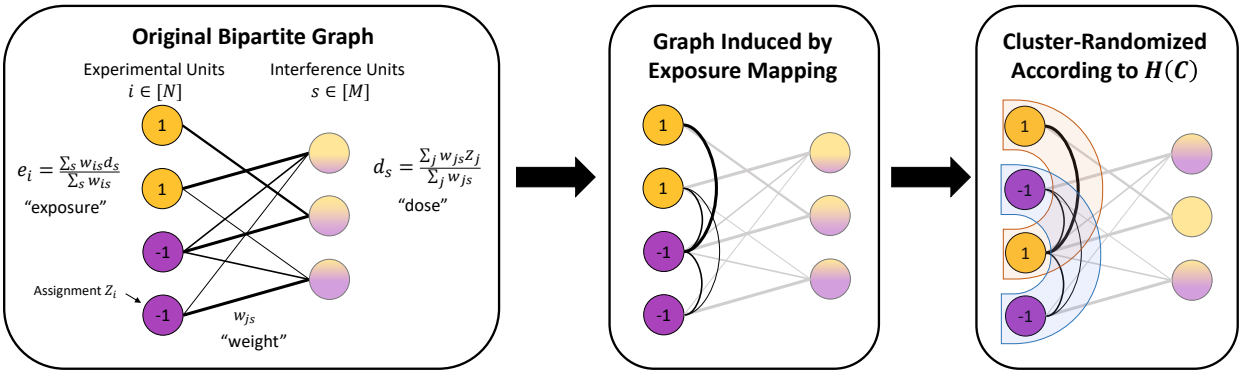


Figure 5.1: Overview of our cluster-randomized experimental design. Left Panel: We are given a bipartite graph connecting experimental units (left) with interference units (right) with edges of known weight w. Assigning treatments $Z_i \in \{-1, 1\}$ to each experimental unit induces a dose d_s on each interference unit, and an exposure e_i on each experimental unit. Middle Panel: The exposure of experimental unit i depends on the assignment of other experimental units j, which induces a graph on the experimental units. Right Panel: Clustering the experimental units based on the induced graph creates clusters of units which heavily influence each others' outcomes. Randomizing the treatment assignment at the level of the cluster provides exposures e_i that are much closer to the treatment Z_i than would be achieved with a unit-randomized design.

exposure mapping is a function $e_i : \{-1, 1\}^N \to \mathbb{R}$ such that $Y_i(Z_i, e_i(\mathbf{Z})) = Y_i(\mathbf{Z})$. In other words, the indirect effect of $\mathbf{Z}_{j \neq i}$ on unit i is captured completely by the exposure $e_i(\mathbf{Z})$. When it is clear from context, we will write e_i instead of $e_i(\mathbf{Z})$ to denote the exposure of unit i. If we could design an experiment such that $e_i = Z_i$ for all units i, then no unit would experience interference. However, in real bipartite graphs it is usually not possible to assign treatments Z_i in a way that eliminates interference entirely.

Given an exposure mapping, we can further posit a model for the effect of assignment Z_i and exposure e_i on the outcome Y_i. We principally consider the linear model

$$Y_i(\mathbf{Z}) = \alpha_i + \beta_i Z_i + \gamma_i e_i, \tag{5.1}$$

which is commonly used in the interference literature (Gui et al., 2015; Harshaw et al., 2022), although other models are discussed in Section 5.4.

Following the tradition of the finite population model, first defined by Neyman (1923), we treat the coefficients α_i, β_i and γ_i as fixed but unknown so that the only randomness in the observation model is due to the choice of treatment assignment $\mathbf{Z}$. This contrasts with alternatives such as a model in which coefficients are drawn from some super-population or a model in which the coefficients are common across i with some additive error term ε in the linear model. Note in particular that the coefficients are not necessarily identical across units i, so that we can never identify the individual α_i, β_i and γ_i. The finite population model avoids making assumptions about

a population from which experimental units are drawn, and for this reason is often preferred in the causal inference literature (Rubin, 1990; Ding et al., 2017).

5.2.2 *The exposure model*

The literature on network interference typically defines the exposure e_i as a function of the treatments of the neighbors of i. For example, the exposure mapping might be the (weighted) fraction of neighbors that are treated, or the count of neighbors that are treated. In the bipartite setting, experimental units are never immediate neighbors. Instead, experimental units interact through their relationships with interference units. Defining an exposure mapping that is analogous to those used in the network setting requires an analogous definition of the "neighborhood" of an experimental unit, as well as the weight of the connections from a unit to each of its neighbors.

We propose a bipartite analogue to the neighborhood-based exposure mapping, composed of two parts: the *dose* $d_s \in [-1, 1]$ of each interference unit $s \in [M]$ represents the weighted average of the treatment assignment Z_i of each experimental unit in the neighborhood of s, while the *exposure* $e_i \in [-1, 1]$ of each experimental unit $i \in [N]$ represents the weighted average of doses among the interference units in the neighborhood of i.

$$d_s = \frac{1}{\sum_{i\in[N]} w_{is}} \sum_{i\in[N]} w_{is} Z_i, \qquad e_i = \frac{1}{\sum_{s\in[M]} w_{is}} \sum_{s\in[M]} w_{is} d_s. \tag{5.2}$$

Because the exposure thus defined can be written as a linear combination of the treatment assignments of the two-hop neighbors of unit i, our definition of exposure can be viewed as an exposure mapping where e_i depends on two-hop neighbors, in which we must impute the effective "assignment" of interference units s from the assignments of their neighbors. Depending on the problem instance, it may be more appropriate to write one or both of these terms as an unnormalized linear combination instead of a convex combination; we show in Section 5.4.3 that all of our results apply to the unnormalized setting as well.

5.2.3 *Estimators*

Estimators for the average total treatment effect τ under network interference typically fall into one of two categories: difference-in-means (DIM) estimators and inverse propensity score (IPS) estimators. The former reports the difference between the mean of N_T treated units and the mean of N_C control units, where we treat N_T and N_C as fixed quantities chosen before treatment randomization occurs. The latter requires a notion of which units are "fully exposed" to treatment or to control, perhaps defined by e_i, and reweighs fully exposed observations by the inverse probability of achieving that state.

$$\widehat{\tau}_{DIM} = \sum_{Z_i=1} \frac{Y_i}{N_T} - \sum_{Z_i=-1} \frac{Y_i}{N_C}, \quad \widehat{\tau}_{IPS} = \frac{1}{N} \sum_{i\in[N]} \frac{Y_i \mathbf{1}\{i \text{ fully treated}\}}{\mathbb{P}(i \text{ fully treated})} - \frac{Y_i \mathbf{1}\{i \text{ fully controlled}\}}{\mathbb{P}(i \text{ fully controlled})}$$

The difference in means estimator can suffer from bias in the presence of interference, because both the treated and control means may be biased estimates for their population quantities $n^{-1}\sum_i Y_i(\mathbf{1})$ and $n^{-1}\sum_i Y_i(-\mathbf{1})$, respectively. The IPS estimator is unbiased as long as a "fully treated" unit does in fact behave like a unit in which $\mathbf{Z} = \mathbf{1}$ (and likewise for control), however, this comes at a cost of high variance if the propensity scores are small.

In our linear model, full exposure only occurs when $e_i = Z_i$, which requires all experimental units in the two-hop neighborhood of i to have the same assignment as i. In many realistic bipartite graphs the chance of full exposure for a given unit is very low, increasing the bias in the IPS estimator to unacceptable levels. Appendix D.1 illustrates the high variance of the IPS estimator for a selection of simulated bipartite graphs and a variety of definitions of "full treatment."

The difference-in-means estimator also has the advantage that it outperforms the IPS estimator in the setting of no interference, $Y_i(\mathbf{Z}) = Y_i(Z_i)$, since in this setting $\widehat{\tau}_{DIM}$ is unbiased and has the same or lower variance than $\widehat{\tau}_{IPS}$ for any definition of full exposure. As a result, the difference in means estimator can be seen as an "optimistic" choice of estimator, which makes better use of the data in the case of no interference while incurring bias in the presence of interference. For the reasons described above, we focus on the difference in means estimator in this chapter.

5.3 Experimental design

Having chosen the linear potential outcomes model of Equation (5.1) with the exposure mapping of Equation (5.2) and the difference-in-means estimator $\widehat{\tau}_{DIM}$, we seek a mechanism for randomly assigning experimental units to treatment or control that achieves low mean squared error (MSE) in recovering the average total treatment effect τ. The MSE can be decomposed into the bias and the variance, where the former is caused by interference and the latter is primarily a function of the number of units randomized to treatment and control. Following the majority of the literature (Ugander et al., 2013; Eckles et al., 2017; Candogan et al., 2021), we consider the class of *balanced cluster-randomized designs*, formalized in Definition 5.3.1.

Definition 5.3.1 (Balanced K-cluster randomized design)**.** *Let $\mathcal{C} = \{C_\ell\}_{\ell=1}^K$ be a partition of the N experimental units into K equally sized clusters. A balanced K-cluster randomized design $\mathcal{D}(\mathcal{C})$ is a distribution over vectors $\mathbf{Z} \in \{-1, 1\}^N$ generated by assigning $Z_i = 1$ for all experimental units i belonging to a set of $K_T \in (0, K)$ clusters chosen uniformly at random among all K clusters.*

If the clustering faithfully captures patterns of interference, cluster randomization can increase the chance that a treated (resp. control) unit is highly exposed to treatment (resp. control), thereby reducing the bias in $\widehat{\tau}_{DIM}$.

We control the variance of the estimator by enforcing a balanced clustering, so that assigning a fixed fraction of clusters to treatment results in a consistent fraction of units assigned to treatment. Since cluster randomization can change the effective number of experimental units (see Section 5.6 and Rolnick et al. (2019, Section 4.2)), the variance of $\widehat{\tau}_{DIM}$ will depend on the number of clusters chosen. In practice, the variance under a given clustering can typically be estimated using historical data, often called an A/A test. Practitioners can use these estimates to control

the variance in light of the anticipated effect sizes. The bias due to interference, however, is not estimable with historical data when all such data is observed under the control condition.

Since the variance can be estimated from historical data and controlled by the number of balanced clusters, we choose to focus on identifying a clustering that minimizes the bias of $\widehat{\tau}_{DIM}$ given a fixed number of balanced clusters. Choosing the correct number of clusters to trade off bias and variance is an important direction for future work.

5.3.1 Existing designs do not work in the one-sided bipartite setting

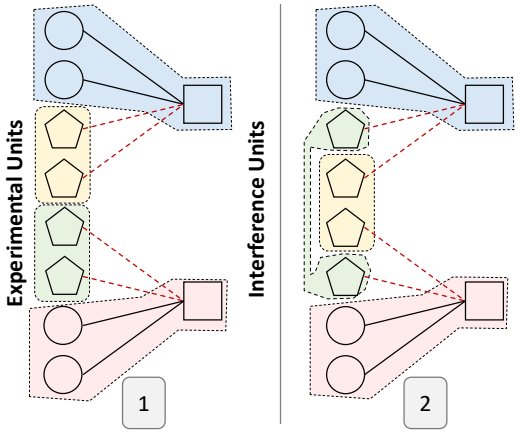
A natural question is whether existing approaches to cluster-randomized designs are adequate for our setting of one-sided experiments on bipartite graphs. In this section, we consider the natural extensions of two clustering algorithms from the literature to our setting and illustrate counterexamples on which each algorithm fails to identify the bias-minimizing clustering.

Direct clustering of the bipartite graph

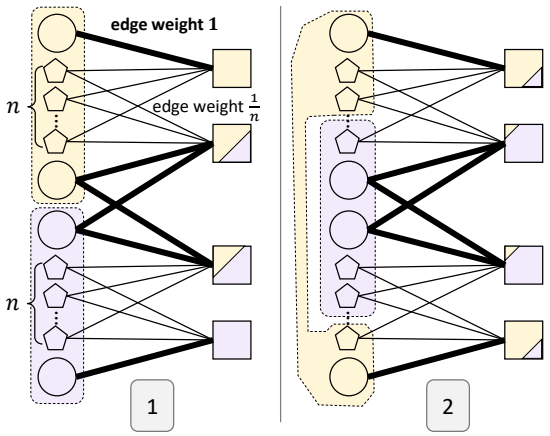
Cluster-randomized designs have primarily been studied in the context of graphs in which all units are experimental units, such as social networks (Ugander et al., 2013; Eckles et al., 2017). Within the class of cluster-randomized designs, balanced partitioning has been suggested as a means to minimize bias (Eckles et al., 2017). The natural analog in the bipartite setting is to create a balanced partitioning of the bipartite graph according to the provided edge weights w_{is}; this extension was considered as a baseline in a different bipartite setting by Pouget-Abadie et al. (2019). Clusters may contain both experimental units and interference units, but only experimental units count toward the balancedness constraint.

Unfortunately, this naive approach fails to take into account the two-hop structure of interference in a bipartite graph. Consider for example two experimental units i and j that must be assigned to a different cluster than that of their common neighbor, interference unit s, perhaps to satisfy a balancedness or cluster cardinality constraint. There is no benefit for the balanced partitioning algorithm on the bipartite graph to assign these two experimental units to the same cluster as each other, despite them sharing a common neighbor.

This mode of failure is illustrated in Figure 5.2a, where a direct clustering of the bipartite graph assigns the same objective function score to the two clusterings, even though the first is bias-optimal under our linear model while the second is not. Both clusterings have cost 4 according to direct partitioning (there are four red dashed lines, representing cut edges). We can use the linearity of $\widehat{\tau}_{DIM}$ to express the bias as a sum of bias contributions of each experimental unit (see Lemma D.2.1 of the appendix). The bias contributed by the circular units is the same in both clusterings, since the expected exposure impurity $|e_i - Z_i|$ is the same in both clusterings and our interference model is linear in e_i and Z_i. However, the exposure of the pentagonal units in Clustering 1 stochastically dominates that of Clustering 2 in the sense of being closer to the true treatment assignment Z_i, which means that Clustering 1 incurs less bias. Therefore, Clustering 1 is superior under our potential outcomes model.



(a) **Failure of the direct balanced partitioning.**



(b) **Failure of maximizing the dose variance.**

Figure 5.2: These counterexamples illustrate the inadequacy of existing cluster-randomized designs in the one-sided bipartite setting. In each case, Clustering 1 is the bias-minimizing clustering under our exposure model (5.2) and linear potential outcomes model (5.1).

Maximizing the variance of the doses

Cluster-randomized designs have also been considered in the setting of two-sided bipartite experiments, in which treatment is assigned to the experimental units but outcomes are measured on the interference units. When the potential outcome of an interference unit is a linear function of the dose d_s, Pouget-Abadie et al. (2019) and Harshaw et al. (2022) recommend assigning treatments $\mathbf{Z}$ in a way that maximizes the empirical variance of the realized doses. In the one-sided bipartite setting it is important to enforce balancedness of the clusters to control the variance of $\widehat{\tau}_{DIM}$, so an extension of their objective to this setting would be $\arg\max_{\mathcal{C}} \mathrm{Tr}(\mathrm{Var}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}(\mathbf{d}))$ over a balanced cluster-randomized assignment, where the dose vector $\mathbf{d}$ is a function of the cluster-randomized assignment $\mathbf{Z} \sim \mathcal{D}(\mathcal{C})$.

Figure 5.2b exhibits a bipartite graph in which the doses d_s are primarily controlled by a small number of experimental units with highly weighted edges, so that maximizing $\mathrm{Tr}(\mathrm{Var}(\mathbf{d}))$ actually reduces the average covariance of Z_i and e_i. Clustering 1 ensures that the majority of experimental units (the $2n$ pentagons) have $e_i = Z_i$ w.p. $\approx \frac{1}{2}$, and $e_i = \frac{3}{4}Z_i$ otherwise, while sacrificing the exposure of the two middle circular units, which have $e_i = 0$ w.p. $\approx \frac{1}{2}$. By contrast, Clustering 2 obtains a high dose variance by prioritizing the assignment of the middle circular units to the same cluster, at the expense of making $e_i = 0$ w.p. $\frac{1}{2}$ for all $2n$ pentagonal units. If we think of the interference among all units as being on the same order of magnitude then Clustering 1, which incurs less interference in the $2n$ pentagonal units, is the correct choice as $n \to \infty$. If, on the other hand, the potential outcomes scale in magnitude with the sum of an experimental unit's

edge weights, then there are settings in which Clustering 2 is bias-optimal. This can be addressed by choosing the appropriate normalization of the dose and exposure for the problem at hand and modifying $\mathcal{H}(\mathcal{C})$ to reflect the choice of normalization; see Section 5.4.3 for further discussion.

5.3.2 The bias-minimizing clustering objective $\mathcal{H}(\mathcal{C})$

Having shown that these two natural extensions fail to identify the bias-minimizing cluster-randomized designs, we propose an alternative clustering objective, $\mathcal{H}(\mathcal{C})$:

$$\mathcal{H}(\mathcal{C}) := \sum_{i\in[N]} \sum_{j\notin\mathcal{C}(i)} \sum_{s\in[M]} \frac{w_{is}}{\sum_s w_{is}} \frac{w_{js}}{\sum_k w_{ks}}.$$

The following lemma shows that our clustering objective is minimax optimal over choices of bounded γ_i in the linear response model of Equation (5.1).

Lemma 5.3.2. *Suppose the difference in means estimator $\widehat{\tau}_{DIM}$ is computed on a balanced K-cluster design $\mathcal{D}(\mathcal{C})$. Let the potential outcomes follow the linear model* (5.1)*. Then the bias is given by:*

$$\tau - \mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}[\widehat{\tau}_{DIM}] = \frac{2}{N} \cdot \frac{K}{K-1} \sum_{i\in[N]} \sum_{j\notin\mathcal{C}(i)} \gamma_i \sum_{s\in[M]} \frac{w_{is}}{\sum_s w_{is}} \frac{w_{js}}{\sum_k w_{ks}}.$$

Suppose further that all interference terms are bounded in magnitude by a constant, so that $\gamma_i = O(1)$. Then the minimax bias is bounded:

$$\begin{aligned}\arg\min_{\mathcal{C}} \max_{\gamma:\ \gamma_i=O(1)} \left|\tau - \mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}[\widehat{\tau}_{DIM}]\right| &= \arg\min_{\mathcal{C}} \sum_{i\in[N]} \sum_{j\notin\mathcal{C}(i)} \sum_{s\in[M]} \frac{w_{is}}{\sum_s w_{is}} \frac{w_{js}}{\sum_k w_{ks}} \\ &= \arg\min_{\mathcal{C}} \mathcal{H}(\mathcal{C}).\end{aligned}$$

To interpret Lemma 5.3.2, we can consider the cases in which the clustering $\mathcal{C}$ can achieve zero bias for $\widehat{\tau}_{DIM}$. One way this can occur is when there is no interference, so that $\gamma_i = 0$ for all i. Even in the presence of interference, the bias of $\widehat{\tau}_{DIM}$ can be zero if the cut edges $w_{is} \ : \ j \notin \mathcal{C}(i)$ are all of zero weight. In such a well-clustered graph, the outcome of unit i depends only on the treatment assignments of other units within its own cluster. Since the same treatment is applied to each element of a cluster, a perfectly clustered graph has $e_i = Z_i$ and experiences no bias due to interference.

Our objective $\mathcal{H}(\mathcal{C})$ has two interpretations: as a clustering objective on an induced graph on experimental units, and as a statistical objective on the covariance between assignments and exposures.

A graphical interpretation of $\mathcal{H}(\mathcal{C})$. Interference occurs when the treatment assignment of unit j affects the exposure of unit i. The influence of Z_j on e_i under the linear interference model (5.1) is

$$e_i(\mathbf{Z}_{j^+}) - e_i(\mathbf{Z}_{j^-}) = 2 \sum_{s\in[M]} \frac{w_{is}}{\sum_s w_{is}} \frac{w_{js}}{\sum_k w_{ks}}. \tag{5.3}$$

where $\mathbf{Z}_{j^+}$ (resp. $\mathbf{Z}_{j^-}$) is the vector $\mathbf{Z}$ with entry Z_j set to 1 (resp. -1). Now consider a directed graph on experimental units where the edge from i to j is weighted according to $e_i(\mathbf{Z}_{j^+}) - e_i(\mathbf{Z}_{j^-})$; clustering according to this graph minimizes the objective $\mathcal{H}(\mathcal{C})$. This is in fact similar to an approach adopted by Rolnick et al. (2019), which studied a similar folding of the graph and showed it to minimize the bias under a linear potential outcome model motivated by geographical migration.

A statistical interpretation of $\mathcal{H}(\mathcal{C})$. A natural goal of a cluster-randomized design is to ensure that the exposure e_i is close to the treatment Z_i. Lemma 5.3.3 shows that our clustering objective $\mathcal{H}(\mathcal{C})$ maximizes the covariance between the exposures and assignments.

Lemma 5.3.3. *Let $\mathcal{D}(\mathcal{C})$ be a balanced K-cluster randomized design. Then we have:*

$$\arg\max_{\mathcal{C}} Tr\left(Cov_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}(\mathbf{Z}, \mathbf{e})\right) = \arg\min_{\mathcal{C}} \mathcal{H}(\mathcal{C}).$$

Finally, we observe that our cluster-randomized design is an instance of the well-known *balanced partitioning* problem (Andreev and Racke, 2006) on the graph over experimental units, with edges given by the explicit formula (5.3). Balanced partitioning is NP-hard, even in its relaxed form which enforces only partial balancedness, but many tools exist to compute it approximately (Andreev and Racke, 2006; Aydin et al., 2019; Nishimura and Ugander, 2013).

5.4 Robustness

So far we have shown that a cluster-randomized design with objective $\mathcal{H}(\mathcal{C})$ minimizes the bias of $\widehat{\tau}_{DIM}$ under the normalized exposure mapping in by Equation (5.2) and the linear potential outcomes model in Equation (5.1). A natural question is: how robust is this design to deviations from this model? In this section we analyze the robustness of the design under alternative potential outcomes models, showing that the design remains minimax optimal for Lipschitz potential outcomes and minimizes an upper bound on the bias for a class of potential outcomes models motivated by the exposure mapping literature. Additionally, there may exist experimental settings where the exposure e_i is better modeled without the normalization constant $\sum_s w_{is}$, such as when the outcome Y_i is proportional to the edge weight incident to unit i. In this case, or in the case that the dose d_s is unnormalized, we show that all of our results remain valid for a generalization of the clustering objective which removes the corresponding normalization constant(s) from $\mathcal{H}(\mathcal{C})$.

5.4.1 Lipschitz potential outcomes

Let $Y_i(Z, e) \in \text{Lip}_L(e)$ be a Lipschitz function in the exposure e, with Lipschitz constant L. Then we can bound the bias in $\widehat{\tau}_{DIM}$ by a multiple of $\mathcal{H}(\mathcal{C})$, and this bound is tight in a minimax sense.

Lemma 5.4.1. *Let observations Y_i be observed from a K-cluster design, with potential outcomes given by $Y_i(Z, e) \in Lip_L(e)$. Then the bias of the difference-in-means estimator $\widehat{\tau}_{DIM}$ is bounded*

above by

$$|\mathbb{E}[\widehat{\tau}_{DIM}] - \tau^*| \leq \frac{2}{N}\frac{K}{K-1}L \cdot \mathcal{H}(\mathcal{C}).$$

Furthermore, this bound is tight over the class of Lipshitz functions, so that balanced clustering according to the objective $\mathcal{H}(\mathcal{C})$ is tight in a minimax sense among all balanced clusterings:

$$\arg\min_{\mathcal{C}} \max_{f \in Lip_L(e)} |\mathbb{E}[\widehat{\tau}_{DIM}] - \tau^*| = \arg\min_{\mathcal{C}} \mathcal{H}(\mathcal{C}).$$

We note that the objective $\mathcal{H}(\mathcal{C})$ is minimax optimal over the class of L-Lipschitz functions regardless of the Lipschitz constant L. As a result, the practitioner need not know this constant in order to find the minimax optimal clustering.

5.4.2 Functions constant in neighborhoods of $\{-1, 1\}$

A common assumption in the exposure mapping literature is that units with exposure e_i close enough to their treatment assignment Z_i behave as if their entire neighborhood were assigned to Z_i. This assumption motivates the use of IPS estimators, which are unbiased in that setting. If we let Δ denote the neighborhood of Z_i in which an exposure is considered fully treated or controlled, then we have the following constraint on the potential outcome function.

$$|Y_i(Z, e) - Y_i(Z, Z)| \begin{cases} = 0 & \text{if } |Z - e| < \Delta \\ \leq B & \text{otherwise} \end{cases} \qquad \forall Z \in \{-1, 1\},\ \forall e \in [-1, 1] \tag{5.4}$$

Under this assumption on the behavior of Y_i, we can upper bound the bias of the difference-in-means estimator by a quantity that will turn out to be minimized by minimizing $\mathcal{H}(\mathcal{C})$.

Lemma 5.4.2. *Let observations Y_i be observed from a K-cluster design, with potential outcomes satisfying Equation* (5.4)*. Then the bias of the difference-in-means estimate $\widehat{\tau}_{DIM}$ is bounded above by*

$$|\mathbb{E}[\widehat{\tau}_{DIM}] - \tau^*| \leq \frac{2B}{N\Delta}\frac{K}{K-1}\mathcal{H}(\mathcal{C}).$$

Lemma 5.4.2 presents only an upper bound on the bias, and in general this bound is not tight. However, we provide simulations in Section 5.5.2 showing that the objective $\mathcal{H}(\mathcal{C})$ is a reasonable heuristic for functions that satisfy Equation (5.4).

5.4.3 Alternative normalizations of the exposure and dose models

Recall that, under the definition of exposure given in Equation (5.2), both the dose d_s and the exposure e_i are normalized by the sum of adjacent edges so that they lie in the range $[-1, 1]$. Depending on the nature of the experiment, it may instead be appropriate to define an exposure mapping in which either or both of these quantities are unnormalized, so that $\tilde{d}_s = \sum_i w_{is} Z_i$

and/or $\tilde{e}_i = \sum_s w_{is} d_s$. For example, suppose that the edge weights w_{is} measure the value of goods purchased by buyer i from seller s and that the treatment results in a multiplicative increase in the cost per unit good. In this setting, a more appropriate model might be one with unnormalized exposure, for example $Y_i = \gamma_i \sum_s w_{is} d_s$ where γ_i represents the multiplicative increase in the cost per unit good for customer i under treatment (the individual treatment effect). This is the approach to normalization taken by Rolnick et al. (2019), who study the bipartite model in the context of search queries issued by users (experimental units) across various geographical regions (interference units). Their potential outcomes model assumes a normalized dose but unnormalized exposure.

We note that we could also write this unnormalized exposure model in terms of the original linear model (5.1) by absorbing the exposure normalization term $(\sum_s w_{is})^{-1}$ into γ_i. However, if the normalization terms varied significantly between buyers i while the individual treatment effects were approximately equal, then the minimax guarantee in Lemma 5.3.2 would be less meaningful when applied to the normalized model than the unnormalized exposure model. In general, the normalization of the dose and exposure should be chosen so that the magnitude of γ_i in the linear model (5.1) has the least variance between experimental units i as possible, as this is the setting in which the minimax result of Lemma 5.3.2 is the tightest.

We can define the generalized objective parameterized by the definitions of dose and exposure (normalized or unnormalized) using the notation of Equation (5.3):

$$\mathcal{H}_{e,d}(\mathcal{C}) = \sum_{i \in [N]} \sum_{j \notin \mathcal{C}(i)} \gamma_i \sum_s \tfrac{1}{2} \left(e_i(\mathbf{d}_{s^+}) - e_i(\mathbf{d}_{s^-}) \right) \cdot \tfrac{1}{2} \left(d_s(\mathbf{Z}_{j^+}) - d_s(\mathbf{Z}_{j^-}) \right)$$

where $\mathbf{d}_{s^+}$ (resp. $\mathbf{d}_{s^-}$) is the vector $\mathbf{d}$ with entry d_s set to 1 (resp. -1), and analogously for $\mathbf{Z}_{j^+}$ and $\mathbf{Z}_{j^-}$. In the setting of normalized dose and response, this simplifies to the original objective $\mathcal{H}(\mathcal{C})$. Choosing the unnormalized dose $\tilde{d}$ yields the difference $\tilde{d}_s(\mathbf{Z}_{j^+}) - \tilde{d}_s(\mathbf{Z}_{j^-}) = 2w_{is}$, and choosing the unnormalized exposure $\tilde{e}$ yields $\tilde{e}_i(\mathbf{d}_{s^+}) - \tilde{e}_i(\mathbf{d}_{s^-}) = 2w_{js}$.

Under the linear potential outcomes model, the generalized objective satisfies the same properties as the original objective function: in particular, it is bias-minimizing in a minimax sense over all $\gamma_i \in [\Gamma_0, \Gamma_1]$, it minimizes $\mathrm{Tr}(\mathrm{Cov}(\mathbf{Z}, \mathbf{e}))$ for the given definition of e, and it has a graphical interpretation as minimizing the cut edges among an induced graph on experimental units. The first two properties can be verified via straightforward modifications of the proofs of Lemmas 5.3.2 and 5.3.3, respectively, while the last property can be seen from the analogue of Equation (5.3) in the unnormalized setting.

5.5 Experiments

We explore the performance of our cluster-based randomized design in several settings using simulated graphs. We compare to the baseline of unit-level randomization as well as cluster-level randomization according to several clustering schemes: the true clustering, balanced partitioning on the original bipartite graph (Section 5.3.1), maximizing the $\mathrm{Tr}(\mathrm{Var}(\mathbf{d}))$ objective (Section 5.3.1), and maximizing the expected empirical variance of the doses as motivated by Pouget-Abadie et al.

Table 5.1: Relative bias of $\widehat{\tau}_{DIM}$ as the bipartite stochastic block model changes (see 5.5.1)

	$p = 0.0$	$p = 0.005$	$p = 0.05$	$p = 0.5$
$\mathcal{H}(\mathcal{C})$	**0.02(±0.06)**	**3.90(±0.06)**	**11.54(±0.05)**	12.98(±0.08)
Tr(Var(d)) objective	**0.01(±0.06)**	**3.84(±0.05)**	**11.49(±0.06)**	12.91(±0.07)
Direct clustering	**0.01(±0.05)**	9.10(±0.13)	12.68(±0.07)	12.95(±0.07)
EXPOSURE-DESIGN	0.33(±0.06)	4.06(±0.06)	11.90(±0.07)	13.00(±0.08)
Unit-level randomization	12.44(±0.08)	12.55(±0.08)	12.76(±0.08)	12.95(±0.08)
True clusters	0.01(±0.06)	3.88(±0.06)	11.58(±0.06)	12.96(±0.06)

(2019) and Harshaw et al. (2022). For the first three clustering objectives, we use code provided by the authors of Aydin et al. (2019) to identify a minimum-cost balanced partitioning of the graphs induced by each objective. For the last objective, which is not a balanced partitioning, we use code provided by the authors of Harshaw et al. (2022) to minimize the EXPOSURE-DESIGN objective, reporting results with the hyperparameter ϕ tuned to minimize mean squared error.

5.5.1 Robustness to different graph structures

We begin by studying the performance of our design as the amount of interference varies. We construct a synthetic graph according to the bipartite stochastic block model with $N = 1,000$ experimental units and $M = 2,000$ interference units. Both sides of the graph are partitioned into 20 equally sized groups with label $i \in [20]$. Experimental and interference units with the same label have an edge of weight 1 with probability 0.5, while units with different labels have an edge of weight 1 with probability p. We experiment with values from $p = 0$ (no interference between clusters) to $p = 0.5$ (the absence of an underlying clustering structure). Potential outcomes were drawn according to the linear model (5.1), with coefficients drawn $\alpha_i \sim \mathcal{N}(0, 1)$, $\beta_i \sim \mathcal{N}(1, 1)$, and $\gamma_i \sim \mathcal{N}(-1, 1)$. All clustering designs used $K = 20$ clusters, with $K_T = 10$ clusters assigned to treatment. Table 5.1 shows the relative bias (defined as $|\mathbb{E}[\widehat{\tau}] - \tau|/\tau$) of $\widehat{\tau}_{DIM}$ under each cluster-randomized design; variance is inconsequential in this setting, so the variance and MSE are reported in the supplementary materials. Uncertainty represents the 95% confidence interval as determined by bootstrapping over 100 random draws of treatment assignment $\mathbf{Z}$, over a single draw of the graph and potential outcomes.

As anticipated, all designs incur lower bias when there is less interference in the graph (i.e., when p is smaller). Our clustering objective $\mathcal{H}(\mathcal{C})$ and the $\text{Tr}(\text{Var}(\mathbf{d}))$ objective perform on par with the true clustering for all values of p. The low bias of the $\mathcal{H}(\mathcal{C})$ objective is unsurprising given the minimax optimality result in Lemma 5.3.2. To understand why the $\text{Tr}(\text{Var}(\mathbf{d}))$ objective behaves so well, we observe that it can be written in a form similar to that of $\mathcal{H}(\mathcal{C})$, but with a

Table 5.2: Relative bias of $\widehat{\tau}_{DIM}$ as the neighborhood of pure exposure, Δ, widens (see 5.5.2)

	$\Delta = 0.1$	$\Delta = 0.3$	$\Delta = 0.5$
$\mathcal{H}(\mathcal{C})$	1.000($\pm$0.004)	**0.457($\pm$0.005)**	**0.001($\pm$0.000)**
Tr(Var(d)) objective	1.002($\pm$0.004)	**0.460($\pm$0.004)**	**0.000($\pm$0.000)**
Direct clustering	0.997($\pm$0.005)	0.950($\pm$0.008)	0.600($\pm$0.020)
EXPOSURE-DESIGN	1.001($\pm$0.004)	0.509($\pm$0.005)	0.009($\pm$0.001)
Unit-level randomization	0.998($\pm$0.004)	1.000($\pm$0.004)	0.998($\pm$0.003)
True clusters	1.001($\pm$0.004)	0.458($\pm$0.004)	0.001($\pm$0.000)

different normalization on w_{is}:

$$\arg\max_{\mathcal{C}} \text{Tr}\left(\text{Var}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}(\mathbf{d})\right) = \arg\min_{\mathcal{C}} \sum_{i\in[N]} \sum_{j\notin\mathcal{C}(i)} \sum_{s\in[M]} \frac{w_{is}}{\sum_k w_{ks}} \frac{w_{js}}{\sum_k w_{ks}}. \tag{5.5}$$

In the special case that the exposure normalization term $\sum_s w_{is}$ is equal to the dose normalization term $\sum_k w_{ks}$ for all i and s, these constants drop out of the $\arg\min$ and the $\text{Tr}(\text{Var}(\mathbf{d}))$ objective becomes equivalent to the $\mathcal{H}(\mathcal{C})$ objective. This is approximately true in the bipartite stochastic block model with equally sized blocks, since all experimental units have the same expected number of connections (and the same for the interference units), which helps to explain why the $\text{Tr}(\text{Var}(\mathbf{d}))$ is so competitive in our experiments.

5.5.2 Robustness to nonlinearity

Next we explore the robustness of our design to nonlinearity in the potential outcomes model Y_i by simulating outcomes according to Equation (5.4), in which the potential outcome is constant for e_i in a Δ neighborhood of Z_i. We simulate $Y_i = -Z_i$ if $|e_i - Z_i| < \Delta$ and $Y_i \sim \mathcal{U}(-1, 1)$ otherwise, encoding a setting in which no knowledge can be gained about τ when e is Δ-far from Z. The graph is given by the bipartite stochastic block model described in Section 5.5.1 with 20 groups, with $p = 0.5$ chance of an edge between units belonging to the same group and $p = 0.005$ for units belonging to different groups. All clustering designs used $K = 20$ clusters, with $K_T = 10$ clusters assigned to treatment. Table 5.2 shows the results of these experiments. Uncertainty represents the 95% confidence interval as determined by bootstrapping over 100 random draws of treatment assignment $\mathbf{Z}$, over a single draw of the graph and potential outcomes.

We see that all designs have lower bias when Δ is large, reflecting the fact that observations with exposure e_i Δ-far from Z_i are useless in determining τ, and it is easier to get Δ-pure observations for large Δ. The results in Table 5.2 support our claim that $\mathcal{H}(\mathcal{C})$ is a reasonable design heuristic for this setting (motivated by the bias upper bound in Lemma 5.4.2). We observe again that the $\text{Tr}(\text{Var}(\mathbf{d}))$ objective does as well as the $\mathcal{H}(\mathcal{C})$ objective; as described in Section

5.5.1, this is due to the fact that these objectives are nearly equivalent under the symmetry of our simulated graphs.

5.5.3 Robustness to exposure mapping: an Airbnb case study

So far we have only considered potential outcomes models where the outcome Y_i is a function of the assignment Z_i and the exposure e_i as defined in Equation (5.2). However, in many real-life settings, an exposure model is only an approximation to the true mechanism of interference.

We test the robustness of our experimental design to misspecification of the exposure mapping by simulating outcomes according to a model developed for vacation rentals by Li et al. (2022), in which there is no explicit exposure mapping defined. We call the experimental units *customers* and the interference units *listings*. In the first phase of the model, each customer i applies to each listing s with probability ϕ_{is}. In the second phase, listings with applications randomly select an application to accept. The measured outcome Y_i is 1 if customer i successfully booked a listing, and 0 otherwise. In alignment with previous work in this literature (Li et al., 2022; Johari et al., 2022) we create a natural clustering structure on the network by assigning each customer and each listing to one of 20 types. Our simulation takes $N = 500$ customers and $M = 1000$ listings, with application probability under the control assignment of $\phi_{is} = 0.016$ if i and s are of the same type, and $\phi_{is} = 0.0001$ otherwise. Treating customer i increases these application probabilities by a factor of α.

When running this marketplace experiment, an experimenter would typically have access to historical data about the rate of successful applications of customer i to listing s, but would not know the consideration probabilities ϕ_{is}. To be faithful to this observation model, we constructed the bipartite graph using twelve rounds of interaction in this marketplace under the control condition. This graph was sampled once and fixed for all experiments. All clustering designs used $K = 20$ clusters, with $K_T = 10$ assigned to treatment. Simulations were performed by extending code provided by the authors of Li et al. (2022) to the cluster-randomized setting. The relative bias, standard deviation, and root mean squared error (RMSE) of $\widehat{\tau}_{DIM}$ of various clustering algorithms are shown in Figure 5.3.

We observe that all clustering algorithms had similar variance, but that the bias of $\mathcal{H}(\mathcal{C})$ and $\mathrm{Tr}(\mathrm{Var}(\mathbf{d}))$ were closest to that of the true clustering, giving those two methods the lowest RMSE. Notably, the unit-randomized design is suggested by Li et al. (2022) in this setting ($N < M$), but our cluster-randomized design outperforms it. We conclude that our clustering method still outperforms other baselines, even in a setting where the potential outcome does not obey the exposure mapping model we hypothesized.

5.6 Future work

We studied a clustering objective, $\mathcal{H}(\mathcal{C})$, that minimizes the bias of the difference-in-means estimator for one-sided bipartite experiments. When the experimental units are well-clusterable, the cluster-randomized design significantly reduces the bias in $\widehat{\tau}_{DIM}$ when compared with unit-level randomization. However, cluster-based randomization may actually reduce the accuracy

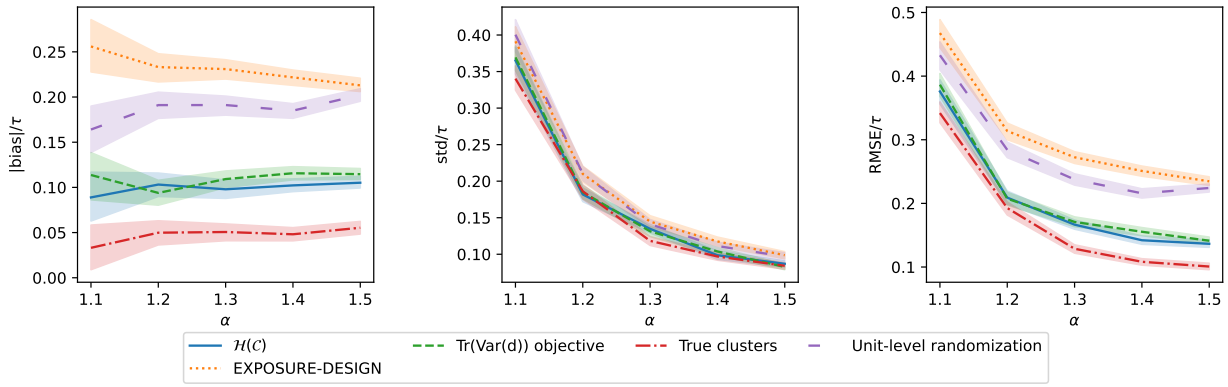


Figure 5.3: Performance of various clustering algorithms as the treatment effect α increases. Uncertainty represents 95% empirical confidence intervals over 500 draws of the treatment assignment $\mathbf{Z}$. Direct balanced partitioning had error 4-8 times the second-worst algorithm, and is not shown.

of the treatment effect estimate if the clustering correlates with individual treatment effects $\tau_i := Y_i(\mathbf{Z} = \mathbf{1}) - Y_i(\mathbf{Z} = -\mathbf{1})$. This may result in significant heterogeneity in the average treatment effect among clusters, which introduces additional variance in the difference-in-means estimator. In the extreme case, the bias-variance tradeoff may favor a unit-randomized design over a cluster-randomized design. We note that this problem is not unique to $\widehat{\tau}_{DIM}$; the IPS estimator also experiences a tradeoff between the variance reduction due to clustering (due to a higher probability of pure exposure) and the variance increase due to heterogeneity between clusters. In practice the variance is typically reduced by choosing a large number of clusters, since the cluster becomes the effective unit of analysis. Of interest is understanding the trade-off incurred by cluster randomization in this setting, and to design techniques to determine, perhaps from historical data, whether cluster randomization should be used for a given experiment.

Appendix A

Appendix for Chapter 2

A.1 Proof of Theorem 2.2.1

In this section, we provide the proof of Theorem 2.2.1.

Proof. We begin with the first part of the theorem, which bounds the probability of overestimating ζ_*. Let A be the event that $\widehat{F}_n$ stays within its Dvoretzky-Kiefer-Wolfowitz (DKW) confidence interval, i.e.,

$$A := \big\{||F_{\nu_*} - \widehat{F}_n||_\infty \leq \tau_{\alpha,n}\big\}.$$

By the DKW inequality Massart (1990), we have $P(A^c) \leq \alpha$. If we assume that event A holds, then

$$\begin{aligned}
\widehat{\zeta}_n(\gamma) &= \max\big\{\zeta \geq 0 : \min_{\nu \in S(\zeta,\gamma)} ||F_\nu - \widehat{F}_n||_\infty > \tau_{\alpha,n}\big\} \\
&\overset{(a)}{\leq} \min\big\{\zeta \geq 0 : \min_{\nu \in S(\zeta,\gamma)} ||F_\nu - \widehat{F}_n||_\infty \leq \tau_{\alpha,n}\big\} \\
&\overset{(b)}{\leq} \zeta^*(\gamma)
\end{aligned}$$

where (a) holds because $S(\zeta,\gamma) \subseteq S(\zeta',\gamma)$ for all $\zeta \leq \zeta'$, and (b) is true because, by event A, $\zeta_*(\gamma)$ is a member of the set. We note that on event A, this argument holds for all γ. We conclude that, with probability at least $1-\alpha$, we have $\widehat{\zeta}_n(\gamma) \leq \zeta^*(\gamma)$ for all γ.

To prove the power statement, let B be as follows

$$B := \big\{||F_{\nu_*} - \widehat{F}_n||_\infty \leq \tau_{\delta,n}\big\}.$$

Suppose that B holds. Then,

$$\begin{aligned}
\widehat{\zeta}_n(\gamma) &= \max\big\{\zeta \geq 0 : \min_{\nu \in S(\zeta,\gamma)} ||F_\nu - \widehat{F}_n||_\infty \geq \tau_{\alpha,n}\big\} \\
&\overset{(a)}{\geq} \max\big\{\zeta \geq 0 : \min_{\nu \in S(\zeta,\gamma)} ||F_\nu - F_{\nu_*}||_\infty - ||\widehat{F}_n - F_{\nu_*}||_\infty \geq \tau_{\alpha,n}\big\} \\
&\overset{(b)}{\geq} \max\big\{\zeta \geq 0 : \min_{\nu \in S(\zeta,\gamma)} ||F_\nu - F_{\nu_*}||_\infty \geq \tau_{\alpha,n} + \tau_{\delta,n}\big\} \\
&\geq \max\big\{\zeta \geq 0 : \min_{\nu \in S(\zeta,\gamma)} ||F_\nu - F_{\nu_*}||_\infty \geq \sqrt{\tfrac{\log(4/\alpha\delta)}{n}}\big\},
\end{aligned}$$

where (a) is the triangle inequality and (b) applies event B. We can rewrite the assumption on n in the theorem statement as

$$\min_{\nu \in S(\zeta_* - \varepsilon, \gamma)} ||F_\nu - F_{\nu_*}||_\infty \geq \sqrt{\frac{\log(4/\alpha\delta)}{n}}$$

which gives us $\widehat{\zeta}_n(\gamma) \geq \zeta_* - \varepsilon$, simultaneously for all γ, whenever B holds. Since B holds with probability at least $1 - \delta$, this completes the proof. □

A.2 Proofs of testing results for mixtures of two Gaussians

A.2.1 Useful Lemmas

In this section of the appendix, we will provide proofs of the results for hypothesis testing - i.e., determining whether there is any mass above 0 - in the two-spike Gaussian setting. This is the simplest setting we consider.

In order to extend our lower bound in Theorem 2.2.1 to the case of two Gaussian spikes, we need to compute the quantity

$$\min_{\nu \in S(0,0)} ||F_\nu - F_{\nu_*}||_\infty$$

for this setting.

Lemma A.2.1. *Let $\nu_* = (1 - \zeta_*)\delta_0 + \zeta_*\delta_{\gamma_*}$, and let S be defined as in Eqn (2.4). Then,*

$$\min_{\nu \in S(0,0)} ||F_\nu - F_{\nu_*}||_\infty \geq \zeta_* \left(\Phi_\sigma(\tfrac{1}{2}\gamma_*) - \Phi_\sigma(-\tfrac{1}{2}\gamma_*)\right)$$

where Φ_σ is the CDF of the normal distribution $\mathcal{N}(0, \sigma^2)$. Furthermore, if $\gamma_* < \sigma$, then

$$\min_{\nu \in S(0,0)} ||F_\nu - F_{\nu_*}||_\infty \geq \frac{23\zeta_*\gamma_*}{24\sigma\sqrt{2\pi}}$$

Proof. First, we lower bound this minimax problem by a minimization at a specific point, $t = \frac{1}{2}\gamma_*$,

$$\begin{aligned}
\min_{\nu \in S(0,0)} ||F_\nu - F_{\nu_*}||_\infty &= \min_{\nu \in S(0,0)} \sup_{t \in \mathbb{R}} |F_\nu(t) - F_{\nu_*}(t)| \\
&\geq \min_{\nu \in S(0,0)} |F_\nu\left(\tfrac{1}{2}\gamma_*\right) - F_{\nu_*}\left(\tfrac{1}{2}\gamma_*\right)| \\
&\geq \min_{\nu \in S(0,0)} F_\nu\left(\tfrac{1}{2}\gamma_*\right) - F_{\nu_*}\left(\tfrac{1}{2}\gamma_*\right)
\end{aligned}$$

Next, we upper bound $F_\nu(\frac{1}{2}\gamma_*)$. By definition,

$$\begin{aligned}
F_\nu(\tfrac{1}{2}\gamma_*) &= \int_{-\infty}^{\infty} \nu(x)\Phi_\sigma(\tfrac{1}{2}\gamma_* - x)dx \\
&\geq \Phi(\tfrac{1}{2}\gamma_*)
\end{aligned}$$

where the inequality follows by the constraint $\nu \in S(0,0)$, so ν can have no mass above 0. We use this bound, as well as the exact value of $F_{\nu_*}\left(\frac{1}{2}\gamma_*\right)$, to lower bound the minimum,

$$\begin{aligned}\min_{\nu \in S(0,0)} ||F_\nu - F_{\nu_*}||_\infty &\geq \Phi_\sigma(\tfrac{1}{2}\gamma_*) - (1-\zeta_*)\Phi_\sigma(\tfrac{1}{2}\gamma_*) - \zeta_*\Phi_\sigma(-\tfrac{1}{2}\gamma_*) \\ &= \zeta_*\left(\Phi_\sigma(\tfrac{1}{2}\gamma_*) - \Phi_\sigma(-\tfrac{1}{2}\gamma_*)\right)\end{aligned}$$

which proves the first claim. For the second claim, when $\gamma_* < \sigma$, we use the second order Taylor series approximation to the Gaussian density. We have

$$\Phi_\sigma(\tfrac{1}{2}\gamma_*) - \Phi_\sigma(-\tfrac{1}{2}\gamma_*) = 2\mathbb{P}(0 \leq X \leq \tfrac{1}{2}\gamma_*) \tag{A.1}$$

for $X \sim \mathcal{N}(0, \sigma^2)$. Taking a quadratic approximation to the normal density yields

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\left(1 + \frac{1}{2\sigma^2}e^{-c^2/2\sigma^2}\left(\frac{c^2}{\sigma^2} - 1\right)x^2\right)$$

for some $c \in [0, \gamma_*]$. This is minimized for $c = 0$, which gives us

$$f(x) \geq \frac{1}{\sigma\sqrt{2\pi}}\left(1 - \frac{1}{2\sigma^2}x^2\right)$$

We can now lower bound (A.1) as follows,

$$\begin{aligned}\mathbb{P}(0 \leq X \leq \tfrac{1}{2}\gamma_*) &= \frac{1}{\sigma\sqrt{2\pi}}\int_0^{\frac{1}{2}\gamma_*} e^{-x^2/2\sigma^2}dx \\ &\geq \frac{1}{\sigma\sqrt{2\pi}}\int_0^{\frac{1}{2}\gamma_*}(1 - \frac{1}{2\sigma^2}x^2)dx \\ &= \frac{1}{\sigma\sqrt{2\pi}}(x - \tfrac{1}{6\sigma^2}x^3)\big|_{x=0}^{x=\frac{1}{2}\gamma_*} \\ &= \frac{1}{\sigma\sqrt{2\pi}}(\tfrac{1}{2}\gamma_* - \tfrac{1}{48\sigma^2}\gamma_*^3) \\ &= \frac{1}{2\sigma\sqrt{2\pi}}\gamma_*\left(1 - \frac{\gamma_*^2}{24\sigma^2}\right)\end{aligned}$$

Since $\gamma_* < \sigma$, this is always positive. We can also use that fact to bound this by

$$\mathbb{P}(0 \leq X \leq \tfrac{1}{2}\gamma_*) \geq \frac{1}{2\sigma\sqrt{2\pi}}\gamma_*\left(\frac{23}{24}\right)$$

so that

$$\Phi_\sigma(\tfrac{1}{2}\gamma_*) - \Phi_\sigma(-\tfrac{1}{2}\gamma_*) \geq \frac{23}{24\sigma\sqrt{2\pi}}\gamma_*$$

which completes the proof. □

Next, we state and prove another lemma that will be useful later. This result bounds the probability that our estimator $\widehat{\zeta}(0)$ returns 0. This is related to the probability of detecting the existence of alternate hypotheses.

Lemma A.2.2. *Let $\nu_* = (1-\zeta_*)\delta_0 + \zeta_*\delta_{\gamma_*}$, and let $\widehat{\zeta}(0)$ be our estimator from Eqn (2.2) evaluated at 0. Then*

$$\mathbb{P}(\widehat{\zeta}(0) = 0) \leq 2\exp\left(-2n\left(\zeta_*\left(\Phi(\tfrac{1}{2}\gamma_*) - \Phi(-\tfrac{1}{2}\gamma_*)\right) - \tau_{\alpha,n}\right)^2\right)$$

Proof. We begin by substituting in the definition of our estimator. Recall the definition of S, from Eqn (2.4).

$$\begin{aligned}
\mathbb{P}\left(\widehat{\zeta}(0) = 0\right) &= \mathbb{P}\left(\max\left(\zeta \geq 0 : \min_{\nu \in S(0,0)} ||\widehat{F}_n - F_\nu||_\infty > \tau_{\alpha,n}\right) = 0\right) \\
&= \mathbb{P}\left(\min_{\nu \in S(0,0)} ||\widehat{F}_n - F_\nu||_\infty \leq \tau_{\alpha,n}\right) \\
&\leq \mathbb{P}\left(\min_{\nu \in S(0,0)} ||F_\nu - F_{\nu_*}||_\infty - ||\widehat{F}_n - F_{\nu_*}||_\infty \leq \tau_{\alpha,n}\right) \\
&= \mathbb{P}\left(||\widehat{F}_n - F_{\nu_*}||_\infty \geq \min_{\nu \in S(0,0)} ||F_\nu - F_{\nu_*}||_\infty - \tau_{\alpha,n}\right)
\end{aligned}$$

Next, we apply the DKW inequality, along with Lemma A.2.1

$$\begin{aligned}
\mathbb{P}\left(\widehat{\zeta}(0) = 0\right) &\leq 2\exp\left(-2n\left(\min_{\nu \in S(0,0)} ||F_\nu - F_{\nu_*}||_\infty - \tau_{\alpha,n}\right)^2\right) \\
&= 2\exp\left(-2n\left(\zeta_*\left(\Phi(\tfrac{1}{2}\gamma_*) - \Phi(-\tfrac{1}{2}\gamma_*)\right) - \tau_{\alpha,n}\right)^2\right)
\end{aligned}$$

This concludes the proof. □

A.2.2 Proof of Corollary 2.2.2

Proof. **Corollary 2.2.2** This corollary follows immediately from Theorem 2.2.1 with $\alpha = \delta$, $\varepsilon = \zeta_{\nu_*}(0) = \zeta_*$ and $\min_{\nu \in S(0,0)} ||F_\nu - F_{\nu_*}||_\infty$ bounded as in Lemma A.2.1. □

A.2.3 Comparison to finite-sample lower bounds

In this section, we prove the statements in Section 2.2.2.

To show that the sample complexity for our hypothesis test matches the lower bound of Jamieson et al. (2016) up to constants, it is helpful to write our sample complexity as a maximum of two quantities.

Lemma A.2.3. *Let $\delta \leq 1$, $\zeta_* \leq 1$, and $\sigma^2 > \gamma_*^2$. Then*

$$\frac{16\sigma^2 \log\left(\frac{2}{\delta}\right)}{\zeta_*^2 \gamma_*^2} = \max\left\{\frac{1}{\zeta_*}, \frac{16\sigma^2 \log\left(\frac{2}{\delta}\right)}{\zeta_*^2 \gamma_*^2}\right\}$$

Proof. We will prove this by showing that the first term of the max is always smaller than the second. We have

$$\begin{aligned}
\frac{16\sigma^2 \log\left(\frac{2}{\delta}\right)}{\zeta_*^2 \gamma_*^2} &> \frac{16 \log\left(\frac{2}{\delta}\right)}{\zeta_*^2} \\
&\geq \frac{16 \log(2)}{\zeta_*^2} \\
&\geq \frac{1}{\zeta_*^2} \\
&\geq \frac{1}{\zeta_*}
\end{aligned}$$

which concludes the proof. □

We see that our test matches the lower bound given in Jamieson et al. (2016) up to constants, as long as δ is bounded away from 1 (so that the first term in the lower bound cannot be arbitrarily small).

A.2.4 Comparison to Asymptotic Lower Bounds

Next, we prove Proposition 2.2.3.

Proof. **Proposition 2.2.3** Let our test be as specified in the proposition. We can bound the sum of the type 1 error (Theorem 2.2.1) and type 2 error (Lemma A.2.2)

$$\begin{aligned}
&\mathbb{P}(\{\text{err}\}) = \mathbb{P}_0(\{\text{reject } H_0\}) + \mathbb{P}_1(\{\text{fail to reject}\}) \\
&\leq \alpha + \exp\left(-2n\left(\max\left(0, n^{-\beta}\left(\Phi(\sqrt{2r\log n}) - \Phi(-\sqrt{2r\log n})\right) - \tau_{\alpha,n}\right)\right)^2\right) \\
&= \alpha + \exp\left(-2n\left(\max\left(0, n^{-\beta}\left(\Phi(\sqrt{2r\log n}) - \Phi(-\sqrt{2r\log n})\right) - n^{-1/2}\sqrt{\tfrac{1}{2}\log(2/\alpha)}\right)\right)^2\right) \\
&= \alpha + \exp\left(-2n\left(\max\left(0, n^{-1/2}\left(n^{-\beta+\frac{1}{2}}\left(\Phi(\sqrt{2r\log n}) - \Phi(-\sqrt{2r\log n})\right) - \sqrt{\tfrac{1}{2}\log(2/\alpha)}\right)\right)\right)^2\right) \\
&= \frac{1}{n} + \exp\left(-2n\left(\max\left(0, n^{-1/2}\left(n^{-\beta+\frac{1}{2}}\left(\Phi(\sqrt{2r\log n}) - \Phi(-\sqrt{2r\log n})\right) - \sqrt{\tfrac{1}{2}\log(2n)}\right)\right)\right)^2\right)
\end{aligned} \tag{A.2}$$

where in the last line we have taken $\alpha = 1/n$. Note that $\Phi(\sqrt{2r\log n}) - \Phi(-\sqrt{2r\log n})$ increases to 1 as $n \to \infty$, provided that $r \neq 0$. We know that $\beta < \frac{1}{2}$. Therefore, for any $r > 0$, there is some n large enough that

$$\begin{aligned} n^{-\beta+\frac{1}{2}}\left(\Phi(\sqrt{2r\log n}) - \Phi(-\sqrt{2r\log n})\right) - \sqrt{\tfrac{1}{2}\log(2n)} &\geq \frac{1}{2}n^{-\beta+\frac{1}{2}} - \sqrt{\tfrac{1}{2}\log(2n)} \\ &\geq 0 \end{aligned}$$

so we can remove the max in the limit of large n, to get

$$\begin{aligned} \mathbb{P}(\{\text{err}\}) &\leq \frac{1}{n} + \exp\left(-2n\left(n^{-1/2}\left(n^{-\beta+\frac{1}{2}}\left(\Phi(\sqrt{2r\log n}) - \Phi(-\sqrt{2r\log n})\right) - \sqrt{\tfrac{1}{2}\log(2n)}\right)\right)^2\right) \\ &= \frac{1}{n} + \exp\left(-2n\cdot n^{-1}\left(n^{-\beta+\frac{1}{2}}\left(\Phi(\sqrt{2r\log n}) - \Phi(-\sqrt{2r\log n})\right) - \sqrt{\tfrac{1}{2}\log(2n)}\right)^2\right) \\ &= \frac{1}{n} + \exp\left(-2\left(n^{-\beta+\frac{1}{2}}\left(\Phi(\sqrt{2r\log n}) - \Phi(-\sqrt{2r\log n})\right) - \sqrt{\tfrac{1}{2}\log(2n)}\right)^2\right) \end{aligned}$$

Now, we take the limit. Since

$$\left(n^{-\beta+\frac{1}{2}}\left(\Phi(\sqrt{2r\log n}) - \Phi(-\sqrt{2r\log n})\right) - \sqrt{\tfrac{1}{2}\log(2n)}\right)^2 = \omega(1)$$

we have that

$$\lim_{n\to\infty} \frac{1}{n} + \exp\left(-2\left(n^{-\beta+\frac{1}{2}}\left(\Phi(\sqrt{2r\log n}) - \Phi(-\sqrt{2r\log n})\right) - \sqrt{\tfrac{1}{2}\log(2n)}\right)^2\right) = 0$$

which concludes the proof. □

A.2.5 Proof of Proposition 2.3.1

Proof. **Proposition 2.3.1** Recall that m denotes the number of genes we will test. By Corollary 2.2.2 with $\alpha = \delta$, we have that

$$m \geq \frac{\log(\frac{2}{\delta})}{\zeta_*^2(\Phi_\sigma(\frac{1}{2}\gamma_*) - \Phi_\sigma(-\frac{1}{2}\gamma_*))^2}$$

implies

$$\mathbb{P}(\zeta_* - \widehat{\zeta}_m > 0) < \delta.$$

We consider two cases. First, if $\gamma_* < \sigma$, then Corollary 2.2.2 states that the sample size must be at least

$$m \geq \frac{16\sigma^2 \log(\frac{2}{\delta})}{\varepsilon^2 \gamma_*^2}$$

to guarantee $\widehat{\zeta}(0) > 0$ with high probability. We can use the relationship $\sigma^2 = \frac{1}{t}$ to get the first requirement of the lemma.

$$\begin{aligned}\gamma_* &\geq \sqrt{\frac{16\log(\frac{2}{\delta})}{\zeta_*^2 mt}} \\ &= \sqrt{\frac{16\log(\frac{2}{\delta})}{\zeta_*^2 B}}\end{aligned}$$

The second requirement comes from the case of large γ_*. When $\gamma_* \geq \sigma$, we can use a table of values of Φ to find

$$\begin{aligned}\Phi_\sigma(\tfrac{1}{2}\gamma_*) - \Phi_\sigma(-\tfrac{1}{2}\gamma_*) &= 1 - 2\Phi_\sigma(-\tfrac{1}{2}\gamma_*) \\ &\geq 1 - 2\Phi_\sigma(-\tfrac{1}{2}\sigma) \\ &= 1 - 2\Phi_1(-\tfrac{1}{2}) \\ &\geq \frac{3}{10}\end{aligned}$$

We see that in this case, as long as

$$m \geq \frac{4\log(\frac{2}{\delta})}{\zeta_*^2} > \frac{200\log(\frac{2}{\delta})}{81\zeta_*^2}$$

(from Corollary 2.2.2), then we have $\widehat{\zeta}(0) > 0$ with high probability. If both of these requirements hold, then we have $\widehat{\zeta}_n(0) > 0$ with high probability, regardless of the value of γ_*. □

Finally, we prove the statements in Remark 2.3.1: that identifying these alternate hypotheses requires at least nB total replicates (while our estimator only takes B total replicates to count them) as long as the budget satisfies $B = O(n)$, and the fraction of alternates is constant.

Proof. **Remark 2.3.1** Identifying the alternates takes at least order $n\gamma_*^{-2}\log(1/\delta)$, even without correcting for multiple testing, because we require at least order $\gamma_*^{-2}\log(1/\delta)$ samples to estimate the mean of a Gaussian within additive error γ_*, with probability $1-\delta$.

Our test, by contrast, takes the larger of

$$B = O\left(\max\left\{\log(1/\delta)\zeta_*^{-2}\gamma_*^{-2}, t\gamma_*^{-2}\right\}\right).$$

If ζ_* is constant, and if t is constant (which occurs when $B = O(n)$ and we allocate replicates equally across all n hypotheses), then our estimator can count half of the alternates using

$$B = O\left(\log(1/\delta)\gamma_*^{-2}\right),$$

which is n times fewer than the number of samples required to identify the discoveries. □

A.3 Proofs of Corollary 2.2.4 and Lemma 2.2.5 (Estimation results for mixtures of two Gaussians)

In this section, we provide proofs of the upper and lower bounds for estimating the amount of mass above the threshold 0 when the data is drawn from a mixture of two Gaussians.

A.3.1 Estimation upper bound (Corollary 2.2.4)

Corollary 2.2.4 is a consequence of Lemma A.3.1, which bounds the minimum ℓ_∞ distance between F_{ν_*} and any CDF with less than $\frac{1}{2}\zeta_*$ mass above zero, and Theorem 2.2.1, which relates this quantity to the sample complexity of estimating the true amount of mass with accuracy $\varepsilon = \frac{1}{2}\zeta_*$. We begin by stating this key lemma, and then we establish a series of technical lemmas necessary for the proof.

The key proof idea is that the quantity bounded in Lemma A.3.1 is the solution to a convex optimization problem, and we can identify the optimal point:

$$\nu_{OPT} = \left(1 - \tfrac{1}{2}\zeta_*\right)\delta_0 + \tfrac{1}{2}\zeta_*\delta_{2\gamma_*}. \tag{A.3}$$

In this section, we follow the conventions of Boyd & Vandenberghe (2004) (p 127) to describe the optimal value and optimal point of an optimization problem. In particular, the *optimal value* of a minimization problem is the minimum value of the objective function over the constraint set, while an *optimal point* is a point ν_{OPT} in the constraint set that achieves the optimal value. We depart from the notation of Boyd & Vandenberghe in one significant respect: we denote the optimal point by ν_{OPT}, and not ν_* (as we will be using the subscript $*$ to denote the reference distribution in our optimization problem).

The first task is to establish the optimality of this solution. We proceed using the standard machinery of convex optimization, as described by Boyd & Vandenberghe (2004): our problem satisfies strong duality, so we exhibit a dual feasible solution with corresponding primal point given by (A.3). Once we have established optimality of this guess, we use several Taylor series approximations to bound the optimal value by a polynomial in γ_* and ζ_*. After establishing this series of technical lemmas, we will prove Lemma A.3.1.

Lemma A.3.1. *Let $X_i \sim \mathcal{N}(\mu_i, 1)$ and $\mu_i \sim \nu_*$, with $\nu_* = (1-\zeta_*)\delta_0 + \zeta_*\delta_{\gamma_*}$. Let $\zeta_* > 0$ and $\gamma_* \in (0, 1]$. Then*

$$\min_{\nu:\mathbb{P}(\mu>0)<\frac{1}{2}\zeta_*} ||F_\nu - F_{\nu_*}||_\infty \geq 0.01\gamma_*^2\zeta_* \tag{A.4}$$

In order to show that our solution (A.3) to the optimization problem (A.4) is optimal, we need to demonstrate certain properties of the subgradient of the objective function. In order to analyze the subgradient of an ℓ_∞ norm, it is first necessary to characterize the value(s) of t for which the supremum in this sup-norm is attained. Our first technical lemma computes these maximizing values of t when ν is our conjectured optimal value, ν_{OPT}.

Lemma A.3.2. *Let $\gamma_* \in (0,1]$ and $\zeta_* \in (0,1]$. Define*

$$\begin{aligned} f(t) &:= \left| \frac{1}{\sqrt{2\pi}} \int_{x=-\infty}^{t} (1-\tfrac{1}{2}\zeta_*)e^{-x^2/2} + \tfrac{1}{2}\zeta_* e^{-(x-2\gamma_*)^2/2} - (1-\zeta_*)e^{-x^2/2} - \zeta_* e^{-(x-\gamma_*)^2/2} dx \right| \\ &:= \left| \tilde{f}(t) \right| \end{aligned}$$

Then

$$\arg\sup_t f(t) = \{t_+, t_-\}$$

where

$$\begin{aligned} t_+ &:= \tfrac{3}{2}\gamma_* + \tfrac{1}{\gamma_*}\log\left(1 - \sqrt{1-e^{-\gamma_*^2}}\right) \\ t_- &:= \tfrac{3}{2}\gamma_* + \tfrac{1}{\gamma_*}\log\left(1 + \sqrt{1-e^{-\gamma_*^2}}\right) \end{aligned}$$

with sign$(\tilde{f}(t_+)) = 1$ and sign$(\tilde{f}(t_-)) = -1$.

Proof. We begin by arguing that we only need to consider the points at which $\tilde{f}'(t) = 0$. To find the argsup, we examine the critical points of $f(t)$. Note that the critical points of $f(t)$ include all critical points of $\tilde{f}(t)$, as well as points at which $f(t) = 0$. Since we are interested in finding the supremum of $f(t)$, and since $f(t) > 0$ for some t (because the argument to the integral is not identically zero), we can discard any critical point at $f(t) = 0$. We conclude that it suffices to examine the critical points of $\tilde{f}(t)$.

We begin by noting that $\lim_{t\to\infty} \tilde{f}(t) = \lim_{t\to-\infty} \tilde{f}(t) = 0$, so the supremum is not found at extreme values of t. This means we only need to inspect the values of t where the derivative $\tilde{f}'(t) = 0$. We compute this derivative,

$$\begin{aligned} \tilde{f}'(t) &= \frac{d}{dt}\frac{1}{\sqrt{2\pi}} \int_{x=-\infty}^{t} (1-\tfrac{1}{2}\zeta_*)e^{-x^2/2} + \tfrac{1}{2}\zeta_* e^{-(x-2\gamma_*)^2} - (1-\zeta_*)e^{-x^2/2} - \zeta_* e^{-(x-\gamma_*)^2/2} dx \\ &= (1-\tfrac{1}{2}\zeta_*)\tfrac{1}{\sqrt{2\pi}}e^{-t^2/2} + \tfrac{1}{2}\zeta_*\tfrac{1}{\sqrt{2\pi}}e^{-(t-2\gamma_*)^2} - (1-\zeta_*)\tfrac{1}{\sqrt{2\pi}}e^{-t^2/2} - \zeta_*\tfrac{1}{\sqrt{2\pi}}e^{-(t-\gamma_*)^2/2} \\ &= \tfrac{1}{2}\zeta_*\tfrac{1}{\sqrt{2\pi}}e^{-t^2/2} + \tfrac{1}{2}\zeta_*\tfrac{1}{\sqrt{2\pi}}e^{-(t-2\gamma_*)^2/2} - \zeta_*\tfrac{1}{\sqrt{2\pi}}e^{-(t-\gamma_*)^2/2} \end{aligned}$$

by the fundamental theorem of calculus. Next, we set the derivative to zero. Due to the specific coefficients in the definition of $\tilde{f}(t)$, the derivative is quadratic in t and can be solved exactly,

$$\begin{aligned} 0 &= -\tfrac{1}{2}\zeta_*\tfrac{1}{\sqrt{2\pi}}e^{-t^2/2} - \tfrac{1}{2}\zeta_*\tfrac{1}{\sqrt{2\pi}}e^{-(t-2\gamma_*)^2/2} + \zeta_*\tfrac{1}{\sqrt{2\pi}}e^{-(t-\gamma_*)^2/2} \\ &= -\tfrac{1}{2}e^{-t^2/2} - \tfrac{1}{2}e^{-(t-2\gamma_*)^2/2} + e^{-(t-\gamma_*)^2/2} \\ &= e^{-\frac{1}{2}t^2}\left(-\tfrac{1}{2} - \tfrac{1}{2}e^{-(-4t\gamma_* + 4\gamma_*^2)/2} + e^{-(-2t\gamma_* + \gamma_*^2)/2}\right) \\ &= -\tfrac{1}{2} - \tfrac{1}{2}e^{-(-4t\gamma_* + 4\gamma_*^2)/2} + e^{-(-2t\gamma_* + \gamma_*^2)/2} \\ &= -\tfrac{1}{2} - \tfrac{1}{2}e^{2t\gamma_* - 2\gamma_*^2} + e^{t\gamma_* - \frac{1}{2}\gamma_*^2} \\ &= -\tfrac{1}{2} - \tfrac{1}{2}\left(e^{t\gamma_*}\right)^2 e^{-2\gamma_*^2} + e^{t\gamma_*}e^{-\frac{1}{2}\gamma_*^2} \end{aligned}$$

Solving with the quadratic equation gives

$$e^{t\gamma_*} = e^{\frac{3}{2}\gamma_*^2} \pm e^{2\gamma_*^2}\sqrt{e^{-\gamma_*^2} - e^{-2\gamma_*^2}}$$
$$t = \tfrac{3}{2}\gamma_* + \tfrac{1}{\gamma_*}\log\left(1 \pm \sqrt{1 - e^{-\gamma_*^2}}\right)$$

We have found the two extreme values of $\tilde{f}(t)$, which we denote

$$t_- = \tfrac{3}{2}\gamma_* + \tfrac{1}{\gamma_*}\log\left(1 + \sqrt{1 - e^{-\gamma_*^2}}\right)$$
$$t_+ = \tfrac{3}{2}\gamma_* + \tfrac{1}{\gamma_*}\log\left(1 - \sqrt{1 - e^{-\gamma_*^2}}\right)$$

It remains to show that they are both suprema; that is, that they attain the same value.

We will show that $\tilde{f}(t_-) = -\tilde{f}(t_+)$. To do this, first define

$$g(x) := \tfrac{1}{2}e^{-x^2/2} + \tfrac{1}{2}e^{-(x-2\gamma_*)^2/2} - e^{-(x-\gamma_*)^2/2}$$

so that

$$\tilde{f}(t) = \zeta_* \frac{1}{\sqrt{2\pi}} \int_{x=-\infty}^{t} g(x)dx$$

The result will follow from three facts: (1) That g is symmetric about $x = \gamma_*$, (2) That $\tilde{f}(\gamma_*) = 0$, and (3) That $\frac{1}{2}(t_- + t_+) = \gamma_*$. We will prove each of these facts, and then use them to show that $\tilde{f}(t_-) = -\tilde{f}(t_+)$.

The function g is symmetric. To show that g is symmetric about $x = \gamma_*$, we will show that $g(\gamma_* + x) = g(\gamma_* - x)$ for all x.

$$\begin{aligned} g(\gamma_* + x) &= -\frac{1}{2}e^{-(\gamma_*+x)^2/2} - \frac{1}{2}e^{-(\gamma_*+x-2\gamma_*)^2/2} + e^{-(\gamma_*+x-\gamma_*)^2/2} \\ &= -\frac{1}{2}e^{-(\gamma_*-x)^2/2} - \frac{1}{2}e^{-(\gamma_*-x-2\gamma_*)^2/2} + e^{-(\gamma_*-x-\gamma_*)^2/2} \\ &= g(\gamma_* - x) \end{aligned}$$

The function $\tilde{f}$ is zero at γ_*. We compute $\tilde{f}(\gamma_*)$. Our first step is a u-substitution $u = x - \gamma_*$

$$\begin{aligned} \tilde{f}(\gamma_*) &= \int_{x=-\infty}^{\gamma_*} -\tfrac{1}{2}e^{-x^2/2} - \tfrac{1}{2}e^{-(x-2\gamma_*)^2/2} + e^{-(x-\gamma_*)^2/2}dx \\ &= \int_{u=-\infty}^{0} -\tfrac{1}{2}e^{-(u+\gamma_*)^2/2} - \tfrac{1}{2}e^{-(u-\gamma_*)^2/2} + e^{-u^2/2}du \\ &= -\tfrac{1}{2}\left(\Phi(\gamma_*) + \Phi(-\gamma_*)\right) + \Phi(0) \end{aligned}$$

Recall that $\Phi(0) = \frac{1}{2}$, and that $\Phi(-t) = 1 - \Phi(t)$ for any t. This gives us

$$\begin{aligned}\tilde{f}(\gamma_*) &= -\tfrac{1}{2}\left(\Phi(\gamma_*) + 1 - \Phi(\gamma_*)\right) + \tfrac{1}{2} \\ &= -\tfrac{1}{2} + \tfrac{1}{2} \\ &= 0\end{aligned}$$

The average of roots t_- and t_+ is γ_*. We will show this fact via direct computation,

$$\begin{aligned}\frac{1}{2}(t_- + t_+) &= \frac{1}{2}\left(3\gamma_* + \tfrac{1}{\gamma_*}\log\left(\left(1 + \sqrt{1 - e^{-\gamma_*^2}}\right)\left(1 - \sqrt{1 - e^{-\gamma_*^2}}\right)\right)\right) \\ &= \frac{1}{2}\left(3\gamma_* + \tfrac{1}{\gamma_*}\log\left(1 - \left(1 - e^{-\gamma_*^2}\right)\right)\right) \\ &= \frac{1}{2}\left(3\gamma_* + \tfrac{1}{\gamma_*}\log\left(e^{-\gamma_*^2}\right)\right) \\ &= \frac{1}{2}\left(3\gamma_* - \gamma_*\right) \\ &= \gamma_*\end{aligned}$$

Conclusion: the two critical points t_- and t_+ are both suprema. We will now show that $\tilde{f}(t_-) = -\tilde{f}(t_+)$. We begin by relating both quantities to $\tilde{f}(\gamma_*)$. Recall that $t_- > \gamma_* > t_+$. We have

$$\begin{aligned}\tilde{f}(t_+) &= \tilde{f}(\gamma_*) - \zeta_* \frac{1}{\sqrt{2\pi}} \int_{t_+}^{\gamma_*} g(x)dx \\ &= -\zeta_* \frac{1}{\sqrt{2\pi}} \int_{t_+}^{\gamma_*} g(x)dx \\ \tilde{f}(t_-) &= \tilde{f}(\gamma_*) + \zeta_* \frac{1}{\sqrt{2\pi}} \int_{\gamma_*}^{t_-} g(x)dx \\ &= \zeta_* \frac{1}{\sqrt{2\pi}} \int_{\gamma_*}^{t_-} g(x)dx\end{aligned}$$

Since $t_- - \gamma_* = \gamma_* - t_+$ and g is symmetric about γ_*, we conclude that

$$\tilde{f}(t_+) = -\tilde{f}(t_-)$$

The last thing we need to show is that $\tilde{f}(t_+)$ is positive (and, consequently, that $\tilde{f}(t_-)$ is negative). This is a direct consequence of the facts that $\tilde{f}(t)$ only has two critical points, that

$\lim_{t\to\infty} \tilde{f}(t) = \lim_{t\to-\infty} \tilde{f}(t) = 0$ (so we know the function crosses zero at most once, at γ_*), and that

$$\begin{aligned}\tilde{f}(0) &= \zeta_* \left(\tfrac{1}{2}\Phi(0) + \tfrac{1}{2}\Phi(2\gamma_*) - \Phi(\gamma_*)\right) \\ &\geq 0\end{aligned}$$

by the concavity of Φ for $x > 0$. Since $0 < \gamma_*$, this tells us that $\tilde{f}$ is positive for all $t < \gamma_*$, including t_+. This completes our proof. □

In order to show the optimality of our conjectured ν_{OPT}, we will write down the KKT conditions for the optimization problem (A.4) and find points that satisfy them. The following lemmas establish certain properties of the dual points. The statements of Lemmas A.3.3 and A.3.4 are motivated by computations in Lemma A.3.5. Readers may wish to skip these two lemmas on their first reading, and review them after understanding how they are used in the argument of Lemma A.3.5.

Lemma A.3.3. *Let $\gamma_* > 0$. Define*

$$\begin{aligned}h(x) &:= \frac{\Phi(t_- - x) - \Phi(t_- - 2\gamma_*)}{\Phi(t_+ - x) - \Phi(t_+ - 2\gamma_*)} \\ t_+ &:= \tfrac{3}{2}\gamma_* + \tfrac{1}{\gamma_*}\log\left(1 - \sqrt{1 - e^{-\gamma_*}}\right) \\ t_- &:= \tfrac{3}{2}\gamma_* + \tfrac{1}{\gamma_*}\log\left(1 + \sqrt{1 - e^{-\gamma_*}}\right)\end{aligned}$$

Then,

$$h(x) \begin{cases} \leq \frac{c}{1-c} & \text{if } x < 2\gamma_* \\ \geq \frac{c}{1-c} & \text{if } x > 2\gamma_* \end{cases}$$

for c defined by

$$\begin{aligned}c &:= \frac{k}{1+k} \\ k &:= \lim_{x\to 2\gamma_*} h(x) = e^{-\frac{1}{2}(t_- - 2\gamma_*)^2 + \frac{1}{2}(t_+ - 2\gamma_*)^2}\end{aligned}$$

We note that this implies $c \in (0, 1)$.

Proof. We will prove this result by first breaking the function h into its numerator and denominator,

$$\begin{aligned}h(x) &= \frac{f(x)}{g(x)} \\ f(x) &:= \Phi(t_- - x) - \Phi(t_- - 2\gamma_*) \\ g(x) &:= \Phi(t_+ - x) - \Phi(t_+ - 2\gamma_*)\end{aligned}$$

We begin by noting that, for both f and g,

$$f(x), g(x) \begin{cases} \geq 0 \text{ if } x < 2\gamma_* \\ = 0 \text{ if } x = 2\gamma_* \\ \leq 0 \text{ if } x > 2\gamma_* \end{cases}$$

which follows from the fact that the normal CDF Φ is strictly increasing.

We will show that $h(x) > k$ when $x > 2\gamma_*$, and that $h(x) < k$ when $x < 2\gamma_*$. This is equivalent to showing that

$$a(x) := f(x) - k \cdot g(x) < 0 \qquad \forall x \neq 2\gamma_* \tag{A.5}$$

(to show the equivalence, recall that $g(x) < 0$ for $x < 2\gamma_*$). In order to show Eqn (A.5), it suffices to show two things: That

$$a(2\gamma_*) = 0 \tag{A.6}$$

and that

$$a'(x) \begin{cases} \leq 0 & \text{if } x > 2\gamma_* \\ \geq 0 & \text{if } x < 2\gamma_* \end{cases} \tag{A.7}$$

Condition (A.6) is satisfied because $f(2\gamma_*) = g(2\gamma_*) = 0$. To show that constraint (A.7) is satisfied, we will take the derivative of $a(x)$,

$$\begin{aligned} a'(x) &= \frac{1}{\sqrt{2\pi}} \left(k e^{-\frac{1}{2}(t_+ - x)^2} - e^{-\frac{1}{2}(t_- - x)^2} \right) && \text{(A.8)} \\ &= \frac{1}{\sqrt{2\pi}} \left(e^{-\frac{1}{2}(t_- - 2\gamma_*)^2 + \frac{1}{2}(t_+ - 2\gamma_*)^2 - \frac{1}{2}(t_+ - x)^2} - e^{-\frac{1}{2}(t_- - x)^2} \right) \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t_- - 2\gamma_*)^2} \left(\frac{e^{-\frac{1}{2}(t_+ - x)^2}}{e^{-\frac{1}{2}(t_+ - 2\gamma_*)^2}} - \frac{e^{-\frac{1}{2}(t_- - x)^2}}{e^{-\frac{1}{2}(t_- - 2\gamma_*)^2}} \right) && \text{(A.9)} \end{aligned}$$

Our goal is to prove (A.7), which only requires information about the sign of $a'(x)$. Since the leading factor in (A.9) is positive, we can ignore it, and it suffices to show that

$$\frac{e^{-\frac{1}{2}(t_+ - x)^2}}{e^{-\frac{1}{2}(t_+ - 2\gamma_*)^2}} - \frac{e^{-\frac{1}{2}(t_- - x)^2}}{e^{-\frac{1}{2}(t_- - 2\gamma_*)^2}} \begin{cases} \leq 0 & \text{if } x > 2\gamma_* \\ \geq 0 & \text{if } x < 2\gamma_* \end{cases}$$

Rearranging the terms of the expression, we see that this is equivalent to showing that

$$b(x) := (t_+ - x)^2 + (t_- - 2\gamma_*)^2 - (t_- - x)^2 - (t_- - 2\gamma_*) \begin{cases} \geq 0 & \text{if } x > 2\gamma_* \\ \leq 0 & \text{if } x < 2\gamma_* \end{cases} \tag{A.10}$$

We have $b(2\gamma_*) = 0$, so it suffices to show that its derivative is nonnegative. We take the derivative with respect to x,

$$\begin{aligned} b'(x) &= -2(t_+ - x) + 2(t_- x) \\ &= 2\left(t_- - t_+\right) \end{aligned}$$

Since $t_- > t_+$ for $\gamma > 0$, we see that $b'(x) > 0$. We conclude that $b(x)$ satisfies (A.10), which implies (A.5). This shows the desired result, with c such that

$$\begin{aligned} \frac{c}{1-c} &= e^{-\frac{1}{2}(t_- - 2\gamma_*)^2 + \frac{1}{2}(t_+ - 2\gamma_*)^2} \\ &= k \end{aligned}$$

Solving for c yields $c = \frac{k}{1+k}$, as claimed. Finally, we see that $\frac{c}{1-c} > 0$, so $c \in (0, 1)$, which completes the proof. □

Lemma A.3.4. *Let $\gamma_* > 0$ and define*

$$\begin{aligned} \tilde{h}(x) &:= \frac{\Phi(t_- - x) - \Phi(t_-)}{\Phi(t_+ - x) - \Phi(t_+)} \\ t_+ &:= \tfrac{3}{2}\gamma_* + \tfrac{1}{\gamma_*}\log\left(1 - \sqrt{1 - e^{-\gamma_*}}\right) \\ t_- &:= \tfrac{3}{2}\gamma_* + \tfrac{1}{\gamma_*}\log\left(1 + \sqrt{1 - e^{-\gamma_*}}\right) \end{aligned}$$

where Φ is the standard normal CDF. Then

$$\tilde{h}(x) \le \frac{c}{1-c} \; \forall x < 0$$

for c defined by

$$\begin{aligned} c &:= \frac{k}{1+k} \\ k &:= \lim_{x \to 2\gamma_*} h(x) = e^{-\frac{1}{2}(t_- - 2\gamma_*)^2 + \frac{1}{2}(t_+ - 2\gamma_*)^2} \end{aligned}$$

Proof. This proof proceeds similarly to the proof of Lemma A.3.3. We begin by rewriting the function $\tilde{h}$,

$$\begin{aligned} \tilde{h}(x) &= \frac{\tilde{f}(x)}{\tilde{g}(x)} \\ \tilde{f}(x) &:= \Phi(t_- - x) - \Phi(t_-) \\ \tilde{g}(x) &:= \Phi(t_+ - x) - \Phi(t_+) \end{aligned}$$

We begin by noting that

$$\tilde{g}(x) \geq 0 \, \forall x < 0$$

which follows because the CDF is an increasing function. Consequently, showing that $\tilde{h}(x) \leq k$ for $x < 0$ is equivalent to showing that

$$\tilde{a}(x) := \tilde{f}(x) - k \cdot \tilde{g}(x) \leq 0 \qquad \forall x < 0$$

In order to prove this inequality, it suffices to show two things: that

$$\tilde{a}(0) = 0$$

and that

$$\tilde{a}'(x) \geq 0 \qquad \forall x < 0$$

Clearly $\tilde{a}(0) = 0$, since $\tilde{f}(0) = 0$ and $\tilde{g}(0) = 0$. To show that the derivative is positive for negative values of x, we begin by taking the derivative,

$$\tilde{a}'(x) = \frac{1}{\sqrt{2\pi}} \left(k e^{-(t_+ - x)^2/2} - e^{-(t_- - x)^2/2} \right)$$

Now, we realize that the previous equation is the same as Equation (A.8), the derivative of $a(x)$ in Lemma A.3.3. We showed in that proof that $a'(x) \geq 0$ for all $x < 2\gamma_*$, so we conclude that $a'(x) \geq 0$ for all $x < 0$. The proof is complete. □

The next lemma is the crux of the proof of Lemma A.3.1. In this lemma, we find the optimal point for the optimization problem (A.4).

Lemma A.3.5. *Let $\zeta_* \in (0, 1]$, $\gamma_* \in (0, 1]$, and*

$$\nu_* = (1 - \zeta_*)\delta_0 + \zeta_* \delta_{\gamma_*}.$$

Define

$$F_\nu(t) = \mathbb{P}_{\mu \sim \nu, X \sim \mathcal{N}(\mu, 1)}(X \leq t),$$

where ν is a probability distribution over $\mathbb{R}$. Then the optimal point for the optimization problem (A.4) *is given by*

$$\nu_{OPT} := \arg \min_{\nu : \mathbb{P}(\mu > 0) < \frac{1}{2}\zeta_*} ||F_\nu - F_{\nu_*}||_\infty = (1 - \tfrac{1}{2}\zeta_*)\delta_0 + \tfrac{1}{2}\zeta_* \delta_{2\gamma_*}$$

Proof. We will prove this statement for the case where ν is a vector in the simplex $\triangle^d$; the continuous case can be recovered by taking the limit as $d \to \infty$.

Let $\nu \in \triangle^d$ be a distribution over a discrete set of points x_i, with $x \in \mathbb{R}^d$, and $0, \gamma_*, 2\gamma_* \in x$. We will prove the conclusion of the lemma using the tools of convex optimization: We will write down the Lagrange dual of problem (A.4), and present a solution that optimizes the dual.

We begin by writing this problem in standard form,

$$\begin{array}{ll} \text{minimize} & f(\nu) \\ \text{subject to} & \sum_i \nu_i = 1 \\ & \nu \succeq 0 \\ & \sum_{i:x_i>0} \nu_i \leq \frac{1}{2}\zeta_* \end{array}$$

where $\nu \in \mathbb{R}^d$ and

$$f(\nu) = \sup_t \left| \frac{1}{\sqrt{2\pi}} \int_{x=-\infty}^{t} \sum_{i=1}^{d} \nu_i e^{-(x-x_i)^2/2} - (1-\zeta_*)e^{-x^2/2} - \zeta_* e^{-(x-\gamma_*)^2/2} dx \right|$$

Since we will be relying on strong duality, we note that our problem satisfies Slater's condition. For example, one interior feasible point for this problem distributes $\frac{1}{4}\zeta_*$ mass equally across entries where $x_i > 0$, and the remaining $1 - \frac{1}{4}\zeta_*$ mass on the remaining entries. Since $\zeta_* > 0$, this is an interior point of the feasible region.

To apply strong duality, we begin by writing the Lagrangian:

$$L(\nu, \lambda_1, \lambda_\zeta, \lambda_0) = f(\nu) + \lambda_1 \left(\sum_i \nu_i - 1 \right) + \lambda_\zeta \left(\sum_{i:x_i>0} \nu_i - \frac{1}{2}\zeta_* \right) - \lambda_0^T \nu$$

Next, we differentiate with respect to ν,

$$\nabla_\nu L(\nu, \lambda_1, \lambda_\zeta, \lambda_0) = \nabla_\nu f(\nu) + \lambda_1 \mathbf{1} + \lambda_\zeta \mathbf{1}_{i:x_i>0} - \lambda_0 \tag{A.11}$$

We begin by noting that the gradient $\nabla_\nu f(\nu)$ is not always defined, but that the subgradient is.[1] In

[1]We will abuse notation slightly, and use ∇ when referring to subgradients.

this case, the KKT conditions say that an optimal point ν_{OPT} must satisfy the following conditions:

$$\nu_{OPT} \geq 0 \tag{A.12}$$

$$\sum_{i:x_i>0} \nu_{OPT}[i] \leq \frac{1}{2}\zeta_* \tag{A.13}$$

$$\sum_{i} \nu_{OPT}[i] = 1 \tag{A.14}$$

$$\lambda_{\zeta,OPT} \geq 0 \tag{A.15}$$

$$\lambda_{0,OPT} \geq 0 \tag{A.16}$$

$$\lambda_{\zeta,OPT}\left(\sum_{i:x_i>0} \nu_{OPT}[i] - \tfrac{1}{2}\zeta_*\right) = 0 \tag{A.17}$$

$$\lambda_{0_i,OPT}\nu_{OPT}[i] = 0 \qquad \forall i \tag{A.18}$$

$$\nabla_\nu L(\nu, \lambda_{1,OPT}, \lambda_{\zeta,OPT}, \lambda_{0,OPT})\big|_{\nu=\nu_{OPT}} \ni 0 \tag{A.19}$$

To show that ν_{OPT} is indeed an optimal primal point, we will present a feasible set of dual variables that, along with ν_{OPT}, satisfy conditions (A.12)-(A.19). By strong duality, this corresponds to an optimal primal point. We use the following ν_{OPT} and dual variables.

$$\nu_{OPT}[i] = \begin{cases} 1 - \frac{1}{2}\zeta_* & \text{if } x_i = 0 \\ \frac{1}{2}\zeta_* & \text{if } x_i = 2\gamma_* \\ 0 & \text{otherwise} \end{cases}$$

$$\lambda_{1,OPT} = (1-c)\Phi(t_-) - c\Phi(t_+)$$

$$\lambda_{\zeta,OPT} = c\left[\Phi(t_+) - \Phi(t_+ - 2\gamma_*)\right] - (1-c)\left[\Phi(t_-) - \Phi(t_- - 2\gamma_*)\right]$$

$$\lambda_{0,OPT}[i] = \begin{cases} c\left[\Phi(t_+ - x_i) - \Phi(t_+)\right] - (1-c)\left[\Phi(t_- - x_i) - \Phi(t_-)\right] & \text{if } x_i < 0 \\ c\left[\Phi(t_+ - x_i) - \Phi(t_+ - 2\gamma_*)\right] - (1-c)\left[\Phi(t_- - x_i) - \Phi(t_- - 2\gamma_*)\right] & \text{if } x_i > 0 \\ 0 & \text{if } x_i = 0 \end{cases}$$

where we define constants depending only on γ_*,

$$c := \frac{k}{1+k}$$

$$k := e^{\frac{1}{2}(t_+ - 2\gamma_*)^2 - \frac{1}{2}(t_- - 2\gamma_*)^2}$$

$$t_+ := \tfrac{3}{2}\gamma_* + \tfrac{1}{\gamma_*}\log\left(1 - \sqrt{1 - e^{-\gamma_*}}\right)$$

$$t_- := \tfrac{3}{2}\gamma_* + \tfrac{1}{\gamma_*}\log\left(1 + \sqrt{1 - e^{-\gamma_*}}\right)$$

It remains to show that these variables satisfy the KKT conditions. We address each category of KKT condition below.

Primal feasibility The primal feasibility conditions (A.12), (A.13) and (A.14) are all clearly satisfied by our choice of ν_{OPT}.

Dual variable nonnegativity Conditions (A.15) and (A.16) require that the dual variables corresponding to inequality constraints are nonnegative. Both conditions follow from Lemma A.3.3.

For condition (A.15), nonnegativity of $\lambda_{\zeta,OPT}$, note that $\lambda_{\zeta,OPT} \geq 0$ is equivalent to

$$\frac{c}{1-c} \geq \frac{\Phi(t_-) - \Phi(t_- - 2\gamma_*)}{\Phi(t_+) - \Phi(t_+ - 2\gamma_*)}$$

(note the equivalence holds because the denominator of the right hand side is positive). By Lemma A.3.3, with $x = 0$, this condition is satisfied.

For condition (A.16), the constraint is clearly satisfied when $x_i = 0$, so it remains to consider the two cases $x_i < 0$ and $x_i > 0$. When $x_i < 0$, the constraint $\lambda_{0,OPT}[i] \geq 0$ is equivalent to

$$\frac{c}{1-c} \geq \frac{\Phi(t_- - x_i) - \Phi(t_-)}{\Phi(t_+ - x_i) - \Phi(t_+)}$$

By Lemma A.3.4, this constraint is satisfied for our chosen value of c.

When $x_i > 0$, the constraint is clearly satisfied for $x_i = 2\gamma_*$ (since $\lambda_{0,OPT} = 0$ in that case). Otherwise, the constraint is equivalent to

$$c\left[\Phi(t_+ - x_i) - \Phi(t_+ - 2\gamma_*)\right] \geq (1-c)\left[\Phi(t_- - x_i) - \Phi(t_- - 2\gamma_*)\right]$$

which implies the following system of inequalities for the value of $\frac{c}{1-c}$

$$\frac{c}{1-c} \geq \frac{\Phi(t_- - x_i) - \Phi(t_- - 2\gamma_*)}{\Phi(t_+ - x_i) - \Phi(t_+ - 2\gamma_*)} \qquad \text{if } x_i < 2\gamma_*$$

and

$$\frac{c}{1-c} \leq \frac{\Phi(t_- - x_i) - \Phi(t_- - 2\gamma_*)}{\Phi(t_+ - x_i) - \Phi(t_+ - 2\gamma_*)} \qquad \text{if } x_i > 2\gamma_*$$

By Lemma A.3.3, these inequalities are satisfied.

Complementary slackness The complementary slackness conditions (A.17) and (A.18) are both satisfied because of the structure of ν_{OPT} and $\lambda_{0,OPT}$. Condition (A.17) is satisfied because

$$\sum_{i:x_i>0} \nu_{OPT} = \frac{1}{2}\zeta_*$$

Condition (A.18) is satisfied because $\nu_{OPT}[i] = 0$ at all but two values of i; at those values (when $x_i = 0$ and $x_i = 2\gamma_*$), we have $\lambda_{0,OPT}[i] = 0$.

Zero is in the subgradient of the Lagrangian The KKT conditions require that zero must be in the subgradient of the Lagrangian, evaluated at the optimal set of variables. Recall that the gradient of the Lagrangian is given by

$$\nabla_\nu L(\nu, \lambda_1, \lambda_\zeta, \lambda_0) = \nabla_\nu f(\nu) + \lambda_1 \mathbf{1} + \lambda_\zeta \mathbf{1}_{i:x_i>0} - \lambda_0$$

We therefore start by computing the subgradient of the objective function f. Recall that, if $f(\nu) = \sup_t f_t(\nu)$ for functions $f_t(\nu)$ indexed by t, and if $\mathcal{I}(\nu) = \{t \in \mathbb{R} \mid f_t(\nu) = f(\nu)\}$ is the set of indices for which the sup is attained, then the subgradient of f contains the convex combination of the subgradients of the "active" functions whose indices are in $\mathcal{I}(\nu)$,

$$\mathbf{conv} \bigcup_{t \in \mathcal{I}(\nu)} \partial f_t(\nu) \subseteq \partial \sup_{t \in \mathbb{R}} f_t(\nu)$$

In our case, we have

$$f_t(\nu) = \left| \frac{1}{\sqrt{2\pi}} \int_{x=-\infty}^{t} \sum_{i=1}^{d} \nu_i e^{-(x-x_i)^2/2} - (1-\zeta_*)e^{-x^2/2} - \zeta_* e^{-(x-\gamma_*)^2/2} dx \right|$$

Let t_* be a value of t where the supremum is attained. We have

$$\begin{aligned} \nabla_\nu f_{t_*}(\nu) &= \nabla_\nu \left| \frac{1}{\sqrt{2\pi}} \int_{x=-\infty}^{t_*} \sum_{i=1}^{d} \nu_i e^{-(x-x_i)^2/2} - (1-\zeta_*)e^{-x^2/2} - \zeta_* e^{-(x-\gamma_*)^2/2} dx \right| \\ &=: \nabla_\nu g(\nu, t_*) \end{aligned}$$

Now, we can compute a subset of the subgradient,

$$\begin{aligned} \nabla_\nu f(\nu) &\supseteq \mathbf{conv} \bigcup_{t_* \in \mathcal{I}(\nu)} \operatorname{sign}(g(\nu, t_*)) \nabla_\nu \frac{1}{\sqrt{2\pi}} \int_{x=-\infty}^{t_*} \sum_{i=1}^{d} \nu_i e^{-(x-x_i)^2/2} - (1-\zeta_*)e^{-x^2/2} - \zeta_* e^{-(x-\gamma_*)^2/2} dx \\ &= \mathbf{conv} \bigcup_{t_* \in \mathcal{I}(\nu)} \operatorname{sign}(g(\nu, t_*)) \Phi(t_* - \mathbf{x}) dx \end{aligned}$$

Next, we evaluate this derivative at $\nu = \nu_{OPT}$. For our choice of ν_{OPT}, Lemma A.3.2 tells us the values of t_*,

$$t_* = \tfrac{3}{2}\gamma_* + \tfrac{1}{\gamma_*} \log\left(1 \pm \sqrt{1 - e^{-\gamma_*^2}}\right)$$

where there are two roots: t_+ sets $\operatorname{sign}(g(\nu_{OPT}, t_+)) = 1$, and t_- sets $\operatorname{sign}(g(\nu_{OPT}, t_-)) = -1$. This implies that the subgradient evaluated at ν_{OPT} contains the convex combination

$$\begin{aligned} \nabla_\nu f(\nu)\big|_{\nu=\nu_{OPT}} &\supseteq p\Phi(\mathbf{1}t_+ - \mathbf{x}) + (1-p)\Phi(\mathbf{1}t_- - \mathbf{x}) \\ \nabla_\nu f(\nu)[i]\big|_{\nu=\nu_{OPT}} &\supseteq p\Phi(t_+ - x_i) + (1-p)\Phi(t_- - x_i) \end{aligned}$$

for some $p \in [0,1]$.

Recall that our goal is to satisfy the KKT condition (A.19), that zero is in the subgradient of the Lagrangian. Having found the subgradient of the objective function, we see this corresponds to showing that

$$\begin{aligned}\nabla_\nu L(\nu, \lambda_{1,OPT}, \lambda_{\zeta,OPT}, \lambda_{0,OPT})|_{\nu=\nu_{OPT}}[i] &\supseteq \begin{cases} p\Phi(t_+ - x_i) + (1-p)\Phi(t_- - x_i) + \lambda_{1,OPT} + \lambda_{\zeta,OPT} - \lambda_{0,OPT}[i] & \text{if } x_i > 0 \\ p\Phi(t_+ - x_i) + (1-p)\Phi(t_- - x_i) + \lambda_{1,OPT} - \lambda_{0,OPT}[i] & \text{if } x_i \leq 0 \end{cases} \\ &= 0\end{aligned}$$

Choosing $p = c$ (which we know is in $[0,1]$ from Lemma A.3.3), and using the values of $\lambda_{1,OPT}$, $\lambda_{\zeta,OPT}$ and $\lambda_{0,OPT}$ we have chosen, we see that this element of the subgradient is in fact zero. This proves that our solution satisfies the final KKT condition.

Conclusion We have proposed a set of primal and dual variables that satisfy the KKT conditions. Since our problem satisfies the conditions of strong duality, we conclude that our choice of primal variables is optimal. □

Now, we are ready to prove Lemma A.3.1.

Proof. (Proof of Lemma A.3.1). We begin by nothing that it suffices to prove this lemma for $\sigma = 1$, since σ is the scale of γ_*.

Lemma A.3.5 gives us the form of ν_{OPT}, which lets us write

$$\begin{aligned}\min_{\nu:\mathbb{P}(\mu>0)<\frac{1}{2}\zeta_*} ||F_\nu - F_{\nu_*}||_\infty &= ||F_{\nu_{OPT}} - F_{\nu_*}||_\infty \\ &= \sup_t |F_{\nu_{OPT}}(t) - F_{\nu_*}(t)|\end{aligned}$$

The supremum over all t is lower bounded by the value at some t. We choose $t = \frac{3}{2}\gamma_* - 1$, which is the first-order Taylor series approximation to t_+ from Lemma A.3.2. This choice lets us bound the quantity below by

$$\begin{aligned}\min_{\nu:\mathbb{P}(\mu>0)<\frac{1}{2}\zeta_*} ||F_\nu - F_{\nu_*}||_\infty &\geq \left| F_{\nu'}\left(\frac{3}{2}\gamma_* - 1\right) - F_{\nu_*}\left(\frac{3}{2}\gamma_* - 1\right)\right| \\ &\geq F_{\nu'}\left(\frac{3}{2}\gamma_* - 1\right) - F_{\nu_*}\left(\frac{3}{2}\gamma_* - 1\right) \\ &= \left(1 - \tfrac{1}{2}\zeta_*\right)\Phi\left(\frac{3}{2}\gamma_* - 1\right) + \tfrac{1}{2}\zeta_*\Phi\left(\frac{3}{2}\gamma_* - 1 - 2\gamma_*\right) \\ &\quad - (1-\zeta_*)\Phi\left(\frac{3}{2}\gamma_* - 1\right) - \zeta_*\Phi\left(\frac{3}{2}\gamma_* - 1 - \gamma_*\right) \\ &= \tfrac{1}{2}\zeta_*\Phi\left(\frac{3}{2}\gamma_* - 1\right) + \tfrac{1}{2}\zeta_*\Phi\left(-\frac{1}{2}\gamma_* - 1\right) - \zeta_*\Phi\left(\frac{1}{2}\gamma_* - 1\right) \\ &= \frac{1}{2}\zeta_*\left(\Phi\left(\frac{3}{2}\gamma_* - 1\right) + \Phi\left(-\frac{1}{2}\gamma_* - 1\right) - 2\Phi\left(\frac{1}{2}\gamma_* - 1\right)\right) \quad \text{(A.20)}\end{aligned}$$

Next, we apply a Taylor series expansion to the normal CDF Φ. We will take this expansion around -1, since we are interested in the behavior for small γ_*. Let $\phi(x)$ be the normal PDF. We have, via a Taylor series,

$$\Phi(x) = \Phi(-1) + \phi(-1)(x+1) + \frac{1}{2}\phi(-1)(x+1)^2 + \frac{1}{6}(x+1)^3\phi(c)(c^2-1)$$

for some $c \in [-1, x]$. We are interested in approximating the CDF at $x = \frac{3}{2}\gamma_* - 1$, so we consider the interval $c \in [-1, -1 + \frac{3}{2}\gamma_*]$. Since $\gamma_* < 1$, we have

$$-\frac{1}{\sqrt{2\pi}} \leq \phi(c)(c^2-1) \leq 0$$

Taylor's remainder theorem lets us compute upper and lower bounds for the CDF on this interval. The upper bound is given by

$$\begin{aligned}
\Phi(x) &\geq \min_{c \in [-1, -1+\frac{3}{2}\gamma_*]} \Phi(-1) + \phi(-1)(x+1) + \frac{1}{2}\phi(-1)(x+1)^2 + \frac{1}{6}(x+1)^3\phi(c)(c^2-1) \\
&\geq \Phi(-1) + \phi(-1)(x+1) + \frac{1}{2}\phi(-1)(x+1)^2 - \frac{1}{6\sqrt{2\pi}}(x+1)^3 \\
&=: \Phi_u(x)
\end{aligned}$$

and the lower bound is

$$\begin{aligned}
\Phi(x) &\leq \max_{c \in [-1, -1+\frac{3}{2}\gamma_*]} \Phi(-1) + \phi(-1)(x+1) + \frac{1}{2}\phi(-1)(x+1)^2 + \frac{1}{6}(x+1)^3\phi(c)(c^2-1) \\
&\leq \Phi(-1) + \phi(-1)(x+1) + \frac{1}{2}\phi(-1)(x+1)^2 \\
&=: \Phi_l(x)
\end{aligned}$$

Applying these bounds to Eqn (A.20), we have

$$\begin{aligned}
\min_{\nu : \mathbb{P}(\mu > 0) < \frac{1}{2}\zeta_*} ||F_\nu - F_{\nu_*}||_\infty &\geq \frac{1}{2}\zeta_* \left(\Phi\left(\frac{3}{2}\gamma_* - 1\right) + \Phi\left(-\frac{1}{2}\gamma_* - 1\right) - 2\Phi\left(\frac{1}{2}\gamma_* - 1\right)\right) \\
&\geq \frac{1}{2}\zeta_* \left(\Phi_l\left(\frac{3}{2}\gamma_* - 1\right) + \Phi_l\left(-\frac{1}{2}\gamma_* - 1\right) - 2\Phi_u\left(\frac{1}{2}\gamma_* - 1\right)\right) \\
&= \frac{1}{2}\phi(-1)\gamma_*^2\zeta_* - \frac{13\gamma_*^3\zeta_*}{48\sqrt{2\pi}}
\end{aligned}$$

We can further simplify this,

$$\begin{aligned}
\min_{\nu : \mathbb{P}(\mu > 0) < \frac{1}{2}\zeta_*} ||F_\nu - F_{\nu_*}||_\infty &\geq \gamma_*^2\zeta_* \left(\frac{1}{2}\phi(-1) - \frac{13\gamma_*}{48\sqrt{2\pi}}\right) \\
&\geq 0.01\gamma_*^2\zeta_*
\end{aligned}$$

which proves the desired result. □

A.3.2 Estimation lower bound (Lemma 2.2.5)

In this section, we prove our finite sample estimation lower bound, Lemma 2.2.5. We begin by stating and proving the main technical lemma we need for this lower bound, a KL divergence calculation for two ε-separated hypotheses. Then, we will prove the lower bound itself, using elements of the standard reduction from estimation to hypothesis testing.

Lemma A.3.6. *Let distributions P_0 and P_1 be mixtures of standard Gaussians defined as*

$$\begin{aligned} P_0 &: (1-\zeta_*)\mathcal{N}(0,1) + \zeta_*\mathcal{N}(\gamma_*, 1) \\ P_1 &: (1-\zeta)\mathcal{N}(0,1) + \zeta\mathcal{N}(\gamma, 1) \end{aligned}$$

where the parameters for P_0 satisfy $\gamma_* \in (0,1)$, $\zeta_* \in (0, \frac{1}{2})$, and the parameters for P_1 are given by

$$\begin{aligned} \zeta &= \zeta_* - \varepsilon \\ \gamma &= \gamma_* \frac{\zeta_*}{\zeta} \end{aligned}$$

so that the free parameters are γ_*, ζ_* and ε. Let $\varepsilon < \frac{2}{3}\zeta_*$. Then the KL divergence between P_0 and P_1 is bounded above by

$$KL(P_1, P_0) \lesssim \varepsilon^2 \gamma_*^4$$

Proof. We begin by bounding the KL divergence by the χ^2 divergence,

$$\begin{aligned} KL(P_1, P_0) &\leq \chi^2(P_1, P_0) \\ &= \int \left(\frac{dP_1 - dP_0}{dP_0}\right)^2 dP_0 \end{aligned}$$

We proceed to bound both the numerator and the denominator of the fraction. If $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$ is the standard normal PDF, then the denominator is bounded by

$$\begin{aligned} dP_0 &= (1-\zeta_*)\phi(t) + \zeta_*\phi(t-\gamma_*) \\ &\geq (1-\zeta_*)\phi(t) \\ &\geq \frac{1}{2}\phi(t) \end{aligned}$$

where we have used the fact that $\zeta_* \in (0, \frac{1}{2})$. The numerator is bounded by

$$\begin{aligned} dP_1 - dP_0 &= (\zeta - \zeta_*)\phi(t) + \zeta_*\phi(t-\gamma_*) - \zeta\phi(t-\gamma) \\ &= \phi(t)\left((\zeta - \zeta_*) + \zeta_* e^{t\gamma_* - \frac{1}{2}\gamma_*^2} - \zeta e^{t\gamma - \frac{1}{2}\gamma^2}\right). \end{aligned}$$

The factor of $\phi(t)$ is now common to both the numerator and the denominator, so they cancel. The KL divergence is now bounded by

$$\begin{aligned} KL(P_1, P_0) &\leq \int 4\left((\zeta - \zeta_*) + \zeta_* e^{t\gamma_* - \frac{1}{2}\gamma_*^2} - \zeta e^{t\gamma - \frac{1}{2}\gamma^2}\right)^2 dP_0 \\ &=: \int 4\Psi(t)^2 dP_0 \\ &= 4\mathbb{E}_0[\Psi(X)^2] \\ &= 4\left((1-\zeta_*)\mathbb{E}_{X\sim\mathcal{N}(0,1)}[\Psi(X)^2] + \zeta_*\mathbb{E}_{X\sim\mathcal{N}(\gamma_*,1)}[\Psi(X)^2]\right) \end{aligned} \tag{A.21}$$

The next step is to bound both of these expectations. To do this, we first expand $\Psi(X)^2$. We have

$$\begin{aligned} \Psi(X)^2 = (\zeta - \zeta_*)^2 &- 2\zeta(\zeta - \zeta_*)e^{\gamma X - \frac{1}{2}\gamma^2} + \zeta^2 e^{2\gamma X - \gamma^2} \\ &+ 2\zeta_*(\zeta - \zeta_*)e^{\gamma_* X - \frac{1}{2}\gamma_*^2} - 2\zeta\zeta_* e^{\gamma X + \gamma_* X - \frac{1}{2}\gamma^2 - \frac{1}{2}\gamma_*^2} + \zeta_*^2 e^{2\gamma_* X - \gamma_*^2} \end{aligned}$$

Note that the random variable X only appears in the form e^{cX}. We will evaluate the two expectations in (A.21) by applying linearity of expectation, and using the moment generating function for a Gaussian random variable. As a reminder, we have

$$\mathbb{E}_{X\sim\mathcal{N}(0,1)}\left[e^{cX}\right] = e^{\frac{1}{2}c^2}$$

and for the shifted Gaussian,

$$\begin{aligned} \mathbb{E}_{X\sim\mathcal{N}(\gamma_*,1)}\left[e^{cX}\right] &= \mathbb{E}_{X\sim\mathcal{N}(\gamma_*,1)}\left[e^{c(X-\gamma_*)}e^{c\gamma_*}\right] \\ &= e^{c\gamma_*}\mathbb{E}_{X\sim\mathcal{N}(\gamma_*,1)}\left[e^{c(X-\gamma_*)}\right] \\ &= e^{c\gamma_*}\mathbb{E}_{X'\sim\mathcal{N}(0,1)}\left[e^{cX'}\right] \\ &= e^{c\gamma_*}e^{\frac{1}{2}c^2} \\ &= e^{c\gamma_* + \frac{1}{2}c^2} \end{aligned}$$

Now, we are ready to evaluate the expectations in (A.21),

$$\begin{aligned} \mathbb{E}_{X\sim\mathcal{N}(0,1)}\left[\Psi(X)^2\right] &= (\zeta - \zeta_*)^2 - 2\zeta(\zeta - \zeta_*) + \zeta^2 e^{\gamma^2} + 2\zeta_*(\zeta - \zeta_*) - 2\zeta\zeta_* e^{\gamma\gamma_*} + \zeta_*^2 e^{\gamma_*^2} \\ &= \zeta_*^2\left(e^{\gamma^2} - 1\right) - 2\zeta\zeta_*\left(e^{\gamma\gamma_*} - 1\right) + \zeta^2\left(e^{\gamma_*^2} - 1\right) \end{aligned} \tag{A.22}$$

and

$$\mathbb{E}_{X\sim\mathcal{N}(\gamma_*,1)}\left[\Psi(X)^2\right] = (\zeta - \zeta_*)^2 - 2\zeta(\zeta - \zeta_*)e^{\gamma\gamma_*} + \zeta^2 e^{\gamma^2 + 2\gamma\gamma_*} + 2\zeta_*(\zeta - \zeta_*)e^{\gamma_*^2} - 2\zeta\zeta_* e^{2\gamma\gamma_* + \gamma_*^2} + \zeta_*^2 e^{3\gamma_*^2}$$

Our next goal is to upper bound both of these expectations, which will allow us to upper bound the KL divergence in (A.21). To bound these expectations, we will use a second order Taylor series

approximation to each exponential term, with a third order remainder term. Recall the expansion of e^x,

$$e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 e^c$$

for some c between 0 and x. Applying this approximation to the exponential terms in (A.22) gives

$$\mathbb{E}_{X\sim\mathcal{N}(0,1)}\left[\Psi(X)^2\right] = \zeta^2\left(\gamma^2 + \frac{1}{2}\gamma^4 + \frac{1}{6}\gamma^6 e^{c_1}\right) - 2\zeta\zeta_*\left(\gamma\gamma_* + \frac{1}{2}\gamma^2\gamma_*^2 + \frac{1}{6}\gamma^3\gamma_*^3 e^{c_2}\right) + \zeta_*^2\left(\gamma_*^2 + \frac{1}{2}\gamma_*^4 + \frac{1}{6}\gamma_*^6 e^{c_3}\right)$$

for $c_1 \in [0, \gamma^2]$, $c_2 \in [0, \gamma\gamma_*]$ and $c_3 \in [0, \gamma_*^2]$. Given our choice of γ, we have $\zeta\gamma = \zeta_*\gamma_*$. Repeated application of this identity lets us simplify the expression above,

$$\begin{aligned}
\mathbb{E}_{X\sim\mathcal{N}(0,1)}\left[\Psi(X)^2\right] &= \frac{1}{2}\left(\zeta\gamma^2 - \zeta_*\gamma_*^2\right)^2 + \frac{1}{6}\zeta^2\gamma^6 e^{c_1} - \frac{1}{3}\zeta\zeta_*\gamma^3\gamma_*^3 e^{c_2} + \frac{1}{6}\zeta_*^2\gamma_*^6 e^{c_3} \\
&\leq \frac{1}{2}\left(\zeta\gamma^2 - \zeta_*\gamma_*^2\right)^2 + \frac{1}{6}e^{\max(c_1,c_2,c_3)}\left(\zeta\gamma^3 - \zeta_*\gamma_*^3\right)^2 \\
&= \frac{1}{2}\left(\zeta\gamma^2 - \zeta_*\gamma_*^2\right)^2 + \frac{1}{6}e^{\gamma^2}\left(\zeta\gamma^3 - \zeta_*\gamma_*^3\right)^2
\end{aligned}$$

where in the last line we have used the fact that $\gamma > \gamma_*$. Our next goal is to show that the second term, which comes from Taylor's remainder theorem, is negligible compared to the first term. To start, note that we can bound γ^2 above by an absolute constant,

$$\begin{aligned}
\gamma^2 &= \gamma_*^2\frac{\zeta_*^2}{\zeta^2} \\
&\leq \frac{\zeta_*^2}{\zeta^2} \\
&\leq 9
\end{aligned}$$

because we required $\zeta_* - \zeta < \frac{2}{3}\zeta_*$, which implies that $\frac{\zeta_*}{\zeta} < 3$. Next, we argue that the second term is of a smaller order than the first,

$$\begin{aligned}
(\zeta\gamma^3 - \zeta_*\gamma_*^3)^2 &= (\zeta_*\gamma^2\gamma_* - \zeta_*\gamma_*^3)^2 \\
&= \zeta_*^2\gamma_*^2\left(\gamma^2 - \gamma_*^2\right)^2 \\
&= \zeta_*^2\gamma_*^2\left((\gamma + \gamma_*)(\gamma - \gamma_*)\right)^2 \\
&= \zeta_*^2\gamma_*^2\left((\gamma_*\frac{\zeta_*}{\zeta} + \gamma_*)(\gamma - \gamma_*)\right)^2 \\
&= \zeta_*^2\gamma_*^2\left(4\gamma_*(\gamma - \gamma_*)\right)^2 \\
&\leq 16\zeta_*^2\gamma_*^2\left(\gamma_*(\gamma - \gamma_*)\right)^2 \\
&= 16\zeta_*^2\gamma_*^2\left(\gamma - \gamma_*\right)^2 \\
&= 16\left(\zeta_*\gamma\gamma_* - \zeta_*\gamma_*^2\right)^2 \\
&= 16\left(\zeta\gamma^2 - \zeta_*\gamma_*^2\right)^2
\end{aligned}$$

We conclude that the expectation can be bounded above in order by just the first term,

$$\begin{aligned}\mathbb{E}_{X\sim\mathcal{N}(0,1)}\left[\Psi(X)^2\right] &\leq c\left(\zeta\gamma^2-\zeta_*\gamma_*^2\right)^2\\ &= c\varepsilon^2\gamma_*^4\left(\frac{\zeta_*}{\zeta_*-\varepsilon}\right)^2\\ &\leq c'\varepsilon^2\gamma_*^4\end{aligned}$$

Next, we argue that the second expectation in (A.21), $\mathbb{E}_{X\sim\mathcal{N}(\gamma_*,1)}[\Psi(X)^2]$, is of the same order. Once we show this, then we can conclude that a linear combination of the two terms is also of that order. We begin, as before, with a Taylor series expansion.

$$\begin{aligned}\mathbb{E}_{X\sim\mathcal{N}(\gamma_*,1)}\left[\Psi(X)^2\right] = (\zeta-\zeta_*)^2 &- 2\zeta(\zeta-\zeta_*)\left(1+\gamma\gamma_*+\frac{1}{2}\gamma^2\gamma_*^2+\frac{1}{6}\gamma^3\gamma_*^3e^{c_1}\right)\\ &+\zeta^2\left(1+\gamma^2+2\gamma\gamma_*+\frac{1}{2}(\gamma^2+2\gamma\gamma_*)^2+\frac{1}{6}(\gamma^2+2\gamma\gamma_*)^3e^{c_2}\right)\\ &+2\zeta_*(\zeta-\zeta_*)\left(1+\gamma_*^2+\frac{1}{2}\gamma_*^4+\frac{1}{6}\gamma_*^6e^{c_3}\right)\\ &-2\zeta\zeta_*\left(1+2\gamma\gamma_*+\gamma_*^2+\frac{1}{2}(2\gamma\gamma_*+\gamma_*^2)^2+\frac{1}{6}(2\gamma\gamma_*+\gamma_*^2)^3e^{c_4}\right)\\ &+\zeta_*^2\left(1+3\gamma_*^2+\frac{9}{2}\gamma_*^4+\frac{27}{6}\gamma_*^6e^{c_5}\right)\end{aligned}$$

The first and second terms from each Taylor series expansion cancel, as a consequence of our choice of γ. The third order terms combine, again through repeated application of the identity $\zeta\gamma=\zeta_*\gamma_*$, to give the following expression,

$$\begin{aligned}\mathbb{E}_{X\sim\mathcal{N}(\gamma_*,1)}\left[\Psi(X)^2\right] = \frac{1}{2}\left(\zeta\gamma^2-\zeta_*\gamma_*^2\right)^2 &- \frac{1}{3}\zeta(\zeta-\zeta_*)\gamma^3\gamma_*^3e^{c_1}+\frac{1}{6}\zeta^2(\gamma^2+2\gamma\gamma_*)^3e^{c_2}\\ &+\frac{1}{3}\zeta_*(\zeta-\zeta_*)\gamma_*^6e^{c_3}-\frac{1}{3}\zeta\zeta_*(2\gamma\gamma_*+\gamma_*^2)^3e^{c_4}+\frac{27}{6}\zeta_*^2\gamma_*^6e^{c_5}\end{aligned}$$

Just like when we bounded the earlier expectation, we note that each of these constants c_i is in fact an absolute constant. Once again, this lets us bound the remainder terms by $c(\zeta\gamma^3-\zeta_*\gamma_*^3)^2$, which we know is smaller in order than the first term above. We conclude that

$$\begin{aligned}\mathbb{E}_{X\sim\mathcal{N}(\gamma_*,1)}\left[\Psi(X)^2\right] &\lesssim \left(\zeta\gamma^2-\zeta_*\gamma_*^2\right)^2\\ &= \varepsilon^2\gamma_*^4\left(\frac{\zeta_*}{\zeta_*-\varepsilon}\right)^2\\ &\leq 9\varepsilon^2\gamma_*^4\end{aligned}$$

Finally, we substitute these bounds on the expectations into our bound for the KL divergence, (A.21),

$$\begin{aligned} KL(P_1, P_0) &\leq 4\left((1-\zeta_*)\mathbb{E}_{X\sim\mathcal{N}(0,1)}[\Psi(X)^2] + \zeta_*\mathbb{E}_{X\sim\mathcal{N}(\gamma_*,1)}[\Psi(X)^2]\right) \\ &\lesssim (1-\zeta_*)\varepsilon^2\gamma_*^4 + \zeta_*\varepsilon^2\gamma_*^4 \\ &\lesssim \varepsilon^2\gamma_*^4 \end{aligned}$$

which completes the proof.

□

Now, we state and prove the lower bound, Lemma 2.2.5.

Proof. Proof of Lemma 2.2.5. We begin by noting that it suffices to prove the lemma for $\sigma = 1$, since σ is the scale of the variable γ_*.

We will prove a minimax lower bound on the number of samples taken by any estimator that estimates ζ within accuracy ε over the set A_ε, with constant probability. We will use a portion of the standard reduction from estimation to hypothesis testing, as can be found in Tsybakov (2009). Specifically, we will prove the statement

$$\inf_{\widehat{\zeta}_n} \sup_{(\zeta,\gamma)\in A_\varepsilon} \mathbb{P}\left(\left|\widehat{\zeta}_n(X) - \zeta\right| \geq \varepsilon\right) \gtrsim e^{-n\varepsilon^2\gamma_*^4}$$

Applying the argument found in Section 2.2 of Tsybakov (2009), along with Theorem 2.2 of the same, we have the bound

$$\inf_{\widehat{\zeta}_n} \sup_{(\zeta,\gamma)\in A_\varepsilon} \mathbb{P}\left(\left|\widehat{\zeta}_n(X) - \zeta\right| \geq \varepsilon\right) \geq \frac{1}{2}e^{-nKL(P_1,P_0)}$$

where P_0 and P_1 are any two parameterizations in A_ε. Choose parameterizations

$$\begin{aligned} \zeta_0 &\in (\zeta_* - 2\varepsilon, \zeta_* + 2\varepsilon) \\ \gamma_0 &\in \left(\tfrac{1}{3}\gamma_*, \tfrac{3}{2}\gamma_*\right) \\ P_0 &= P(\zeta_0, \gamma_0) \\ P_1 &= P\left(\zeta_0 - \varepsilon, \gamma_0\frac{\zeta_0}{\zeta_0 - \varepsilon}\right) \end{aligned}$$

We note that both P_0 and P_1 are in A_ε, due to the constraint $\varepsilon \in (0, \frac{2}{3}\zeta_*)$. Furthermore, by Lemma A.3.6, we have a bound on the KL divergence between P_0 and P_1. We substitute this into our minimax bound,

$$\inf_{\widehat{\zeta}_n} \sup_{(\zeta,\gamma)\in A_\varepsilon} \mathbb{P}\left(\left|\widehat{\zeta}_n(X) - \zeta\right| \geq \varepsilon\right) \gtrsim \varepsilon e^{-n\varepsilon^2\gamma_0^4}$$

This bound holds for any choice of (ζ_0, γ_0) in the ranges described above. But note that, for every (ζ_0, γ_0) in this range, the bound is of the same order, specifically

$$\inf_{\widehat{\zeta}_n} \sup_{(\zeta,\gamma)\in A_\varepsilon} \mathbb{P}\left(\left|\widehat{\zeta}_n(X) - \zeta\right| \geq \varepsilon\right) \gtrsim \varepsilon e^{-n\varepsilon^2\gamma_*^4}.$$

We conclude that any estimator claiming, with constant probability, to estimate ζ with accuracy better than ε over A_ε, and in particular on the instance (ζ_*, γ_*), must take at least

$$n \gtrsim \frac{1}{\varepsilon^2\gamma_*^4}$$

samples. □

A.4 Experimental details and algorithm implementation

A.4.1 Implementation

We implemented our estimator in Python. Instead of directly optimizing Eqn (2.2), which we found lacked robustness, we determined the value of $\widehat{\zeta}_n(\gamma)$ via binary search on the unit interval. The algorithm is shown below. At each stage of binary search, the algorithm performs a hypothesis test to decide whether there is ζ mass above γ. The hypothesis test is identical to the constraint in Eqn (2.2) and the set $S(\zeta, \gamma) \subseteq S(\zeta', \gamma)$ for $\zeta \leq \zeta'$, so this method yields the same results as direct optimization. The optimization was solved using CVXPY Diamond and Boyd (2016); Agrawal et al. (2018), with the ECOS solver and default parameters.

A.4.2 Code availability and computing infrastructure

Code implementing our estimator in Python, including a Jupyter notebook to reproduce Figure 2.3, can be found at `https://github.com/jenniferbrennan/CountingDiscoveries`. We also provide the data from Hao et. al (2008) as a tab-delimited file, to facilitate experiments on their data. Please see the associated README file for an explanation of the data, and an example of loading the data into Python. Experiments were run on an Ubuntu server with 56 cores and 64 GB of RAM.

A.4.3 Experimental details for Poisson and Binomial experiments

In the Poisson experiment, we took $\mu = 1$ as the null hypothesis and drew $n = 100,000$ examples with mean parameters $\lambda_i \sim 0.8\delta_1 + 0.2(\beta(a = 2, b = 5) * 5 + 2)$ (i.e., the alternates means were from a scaled and shifted Beta distribution). In the binomial experiment, we took $n = 100,000$ examples with means drawn from $0.9\delta_{0.5} + 0.1(\beta(a = 2, b = 5) * 0.5 + 0.5)$ and generated test statistics with $t = 20$ trials per binomial random variable. Given that $\mathbb{P}(X_i = 20) = 9 \cdot 10^{-7}$, while the Bonferroni-adjusted critical value for a test at the 0.05 level is $5 \cdot 10^{-7}$, none of the alternate hypotheses in the binomial could be rejected under a FWER guarantee.

Algorithm 1 Binary search to return $\widehat{\zeta}_n$

Input: Confidence level α, n samples $\{X_i\}_{i=1}^n$, and threshold γ
Result: $\widehat{\zeta}_n$, a lower bound estimate on the number of discoveries above threshold γ
Initialize $i_{min} = 0$, $i_{max} = n$, $\tau_{\alpha,n} = \sqrt{\frac{\log(2/\alpha)}{2n}}$
while $i_{max} - i_{min} > 1$ **do**
  $i_{avg} = \lfloor \frac{i_{min}+i_{max}}{2} \rfloor$
  $\zeta = i_{avg}/n$
  Compute test statistic $T(X;\zeta,\gamma) = \min_{\nu \in S(\zeta,\gamma)} ||\widehat{F}_n - F_\nu||_\infty$
  if $T(X;\zeta,\gamma) > \tau_{\alpha,n}$ **then**
    // Reject the null hypothesis; conclude there is at least ζ mass above γ
    $i_{min} = i_{avg}$
  else
    $i_{max} = i_{avg}$
  end if
end while
$\widehat{\zeta}_n = i_{min}/n$
return $\widehat{\zeta}_n$

Appendix B

Appendix for Chapter 3

B.1 Proof of Theorem 3.3.2

We begin by rewriting Definition 3.3.1 in an easier to manipulate form.

Proposition B.1.1. *Suppose the set of drugs Ω does not display paradoxical growth, as defined in Definition 3.3.1, and let $\mathbf{x_0}, \mathbf{x_1} \in \mathcal{X}(\Omega)$ be two antibiotic combinations of drugs in Ω at specified concentrations. If $\mathbf{x_1} \geq \mathbf{x_0}$ coordinatewise, and both $\mathbf{x_0}$ and $\mathbf{x_1}$ display a response above some threshold t (so that $r(\mathbf{x_1}) \geq t$ and $r(\mathbf{x_0}) \geq t$), then any convex combination of the two points also has response above the threshold:*

$$r(\lambda \mathbf{x_0} + (1-\lambda)\mathbf{x_1}) \geq t \; \forall \lambda \in (0,1). \tag{B.1}$$

Proof. We begin by defining $\mathbf{x_\Delta} =: \mathbf{x_1} - \mathbf{x_0} \in \mathbb{R}^{|\Omega|}_{\geq 0}$, which itself can be viewed as its own antibiotic combination. Then, we have

$$\begin{aligned} r(\lambda \mathbf{x_0} + (1-\lambda)\mathbf{x_1}) &= r(\lambda \mathbf{x_0} + (1-\lambda)(\mathbf{x_\Delta} + \mathbf{x_0})) && \text{(B.2)} \\ &= r(\mathbf{x_0} + (1-\lambda)\mathbf{x_\Delta}) && \text{(B.3)} \\ &= r(\mathbf{x_0} + c\mathbf{x_\Delta}) && \text{(B.4)} \end{aligned}$$

for $c \in (0,1)$. Next, we have a case analysis on the relationship between $r(\mathbf{x_0} + c\mathbf{x_\Delta})$ and $r(\mathbf{x_1})$.

Case 1: $r(\mathbf{x_0} + c\mathbf{x_\Delta}) \geq r(\mathbf{x_1})$. In this case, we apply $r(\mathbf{x_1}) \geq t$ to conclude that $r(\lambda \mathbf{x_0} + (1-\lambda)\mathbf{x_1}) = r(\mathbf{x_0} + c\mathbf{x_\Delta}) \geq t$, as desired.

Case 2: $r(\mathbf{x_0} + c\mathbf{x_\Delta}) < r(\mathbf{x_1})$. In this case, we begin with the contrapositive of Assumption 3.3.1 with $\mathbf{x} = \mathbf{x_\Delta}$, $c_2 = c \in (0,1)$, $c_3 = 1$ and $c_1 = 0$. With these assignments, the contrapositive is

$$r(\mathbf{x_1}) > r(\mathbf{x_0} + c\mathbf{x_\Delta}) \implies r(\mathbf{x_0} + c\mathbf{x_\Delta}) \geq r(\mathbf{x_0}). \tag{B.5}$$

The hypothesis of the implication is satisfied, so we conclude that

$$r(\mathbf{x_0} + c\mathbf{x_\Delta}) \geq r(\mathbf{x_0}) \geq t \tag{B.6}$$

as desired. □

With this proposition in hand, we are now ready to prove our main theorem:

Proof of Theorem 3.3.2. Let $\{c_j\}_{j=1}^m$ denote the N_i-normalized concentrations tested under the NDS design. Let c_{j^*} denote the MECI identified by the design, with $j^* \in \{1, \dots, m\}$.

To prove the theorem, we will leverage Proposition B.1.1 to show that Assumption 3.3.1 implies that all combinations of the $d = |\Omega|$ drugs at concentrations with MEC less than c_{j^*} are ineffective, and therefore c_{j^*} is also the MECI among all m^d combinations. Note that, if the MECI corresponds to the smallest sampled concentration $j^* = 1$, we are done, because there is only one combination with MEC c_1, and the NDS design samples it. Therefore, in what follows, we will assume $j^* > 1$.

We will now show that the hypercube of side length c_{j^*-1}, extending from the origin to $c_{j^*-1}\mathbf{1}$, contains only ineffective points. To see this, first note that, under normalized diagonal sampling, all the vertices of this hypercube have been sampled. Next, observe that the vertices must all have come up "ineffective," or else c_{j^*} would not be the minimum-norm effective point. It remains to show that our assumption implies that all other points in the hypercube are ineffective.

By Assumption 3.3.1 and Proposition B.1.1, we know that if $\mathbf{x_1} \geq \mathbf{x_2}$ componentwise and both are ineffective, then $\lambda\mathbf{x_1} + (1-\lambda)\mathbf{x_2}$ is also ineffective for all $\lambda \in [0, 1]$. We will apply this rule recursively d times, to show that the entire hypercube is ineffective.

For the first step, we show that the one-dimensional edges of the hypercube are ineffective, by noting that every point on an edge is a convex combination of two points on the vertices of the cube, and therefore is ineffective. Next, we take the convex combination of all vertices and edges, which shows that the two-dimensional faces are all ineffective. We continue with this argument d times, concluding that all points with MEC less than c_{j^*} must be ineffective. This confirms that the MECI among all m^d possible combinations is the same as the MECI found with the diagonal sampling scheme. □

B.2 Connection between the Minimax Effective Concentration Index and the Highest Single Agent model

In this section, we justify our claim that our MECI gives a dose-effect version of the Highest Single Agent model. To see this it is worth taking an isobologram perspective (Foucquier and Guedj, 2015; Berenbaum, 1977). For concreteness focus on the case of two drugs A, B, and a fixed effect of interest, for example the OD10h. Assume we are given a specific value of the response E, let $x_A^*(E), x_B^*(E)$ be the minimal doses which achieve this value for each drug individually. Dose-effect models describe the set of points in dose space which should also have an effect E. The MECI naturally gives rise to a null model that posits that the corresponding isobole for a combination of drugs with effect E is the set $\{(x_A, x_B) | \max\{\frac{x_A}{x_A^*(E)}, \frac{x_B}{x_B^*(E)}\} = 1\}$. This is in contrast to the Loewe model considers the isobole $\frac{x_A}{x_A^*} + \frac{x_B}{x_B^*} = 1$.

Though dose models do not directly specify the effect of a combination of drugs directly, the effect can be imputed from knowledge of all possible isoboles. More precisely, given the set of isoboles corresponding to the entire range of possible effects E, we can determine for any given concentration (x_A, x_B) what the effect should be - namely the corresponding of the isobole passing through that point. Concretely for the Loewe model to find the expected effect at (x_A, x_B) we would need to find x'_A and x'_B both with the same effect so that the line $x_A/x'_A + x_B/x'_B = 1$ passes

through x_A, x_B. In absence of dose responses for every possible concentration of A, B it is unclear how to find this quantity. In the case of the MECI, it is sufficient to know the effect of x_A and x_B individually to determine the isobole. Namely the effect at any point is just $\max\{r(x_A), r(x_B)\}$. This is effectively the HSA model on effects(Foucquier and Guedj, 2015). We note however that the MECI is different from the CI you would get by using the HSA effect model(Foucquier and Guedj, 2015) since it is defined on concentrations relative to x_A^*, x_B^* not on the effect directly.

B.3 Evidence of non-paradoxical growth

Central to our results is the absence of paradoxical growth, as given in Definition 3.3.1. Exhaustively verifying this assumption for all background combinations $\mathbf{x}_0$ and additive combination $\mathbf{x}$ would be as difficult as exhaustively searching the entire d-dimensional drug space, which is experimentally intractable. However, we can spot-check this assumption by measuring dose-response curves for various pairs of $(\mathbf{x}_0, \mathbf{x})$ and verifying that we observe no paradoxical growth.

We tested 100 randomly chosen pairs of $(\mathbf{x}_0, \mathbf{x})$, each representing a different (off-diagonal) ray in high-dimensional antibiotic space. The 100 pairs we tested represent less than 0.1% of the possible $(\mathbf{x}_0, \mathbf{x})$ combinations; testing more pairs would add confidence to our belief in non-paradoxical growth, but at the expense of additional experimental effort and resources. To select the background concentration vector $\mathbf{x}_0$, we first chose to include or exclude each of the 8 drugs with probability $\frac{1}{2}$. The concentration of included antibiotics was then selected as $\frac{1}{4}$ the MIC of those drugs in combination, as measured during our MIC-normalized experiment. This ensured that the background concentration was itself "ineffective," so that the dose-response curve of $\mathbf{x}_0 + c\mathbf{x}$ gave a nontrivial dose-response curve. Next, we randomly selected a concentration vector $\mathbf{x}$. Included antibiotics were then randomly assigned to be included at either their MIC, $\frac{1}{2}$MIC, or $\frac{1}{4}$MIC. This ensured that we were measuring "off-diagonal" rays in high-dimensional space, which is where we might be concerned about finding paradoxical growth. Note that a drug could be included in both the background set $\mathbf{x}_0$ and the additive set $\mathbf{x}$.

For each choice $(\mathbf{x}_0, \mathbf{x})$, we measured eight concentrations, from six 2-fold steps below the initially chosen concentration of $\mathbf{x}$ to two 2-fold steps above. Measurements were taken in three replicates.

Figure B.1 shows the 100 dose-response curves generated from these experiments. Curves were generated by taking the average of three replicates at each concentration, then smoothed using a Gaussian kernel (bandwidth $\sigma = 1$) to reduce noise (this smoothing operator is less aggressive than the typical 3-point moving average used in, e.g., Katzir et al. (Katzir et al., 2019)). We observe that, among the $(\mathbf{x}_0, \mathbf{x})$ tested, none display paradoxical growth (where paradoxical growth would be evidenced by a curve with AUGC that decreases then increases as the concentration increases, as in panel C of Figure 3.1). This evidence supports the assumption of non-paradoxical growth for our set of antibiotics.

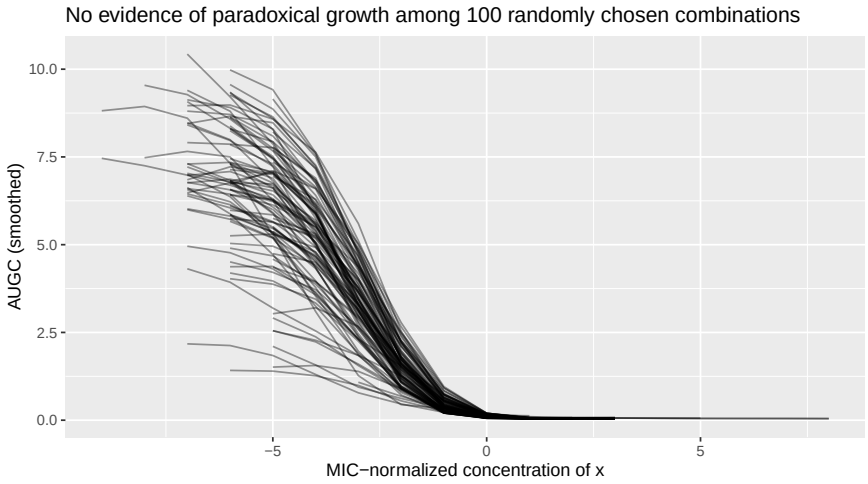


Figure B.1: Dose response curves for 100 randomly chosen baseline/additive pairs $(\mathbf{x}_0, \mathbf{x})$. Curves are translated to reach their MIC at the same point on the graph, to aid legibility. We observe no paradoxical growth; in particular, no curve shows a pattern of decreasing then increasing AUGC as the concentration increases.

B.4 Loewe analysis

In this section we provide a Loewe analysis of our experimental data, for both the breakpoint-normalized and MIC-normalized diagonal designs. To compute the Loewe score for an N drug combination at concentrations $x_1, \ldots, x_N$, we compute the sum of the individual MIC-normalized concentrations:

$$\sum_{i=1}^{N} \frac{x_i}{\text{MIC}_i} \tag{B.7}$$

where MIC_i is the MIC of drug i alone. Figure B.2 shows the distribution of Loewe scores for the breakpoint- and MIC-normalized experiments. Interpretation of synergy and antagonism follow the guidelines of Chou Chou (2006) (Table 4). Note that antagonism is common while synergy is rare, and no combination attained "strong" or "very strong" synergy.

We emphasize that this only captures Loewe synergies for the combinations at the tested concentration ratios; it is possible that combining the antibiotics at different ratios would result in different Loewe scores. This concept is illustrated in Figure B.3, in which the choice of the diagonal on which to sample changes the determination of Loewe synergy.

Figure B.4 shows the Loewe interaction of drug combinations that exhibited "weak synergy" according to the Emergent Synergy Score, with the Loewe interaction determined along the concentration ratio sampled by the NDS design. Observe that the same combination may have a different Loewe synergy score across the two charts because it was measured at a different ratio

in each experiment; this concept is illustrated in Figure B.3.

We see that most combinations that our method finds to be weakly synergistic are either synergistic or additive under the Loewe model, with only three combinations exhibiting Loewe antagonism. The Loewe perspective may also clarify why so many more combinations appeared to be synergistic under the MIC normalization: over 40% (18/44) of these weakly synergistic combinations were drug pairs that were effective when combined at half of their MIC. Such a combination is completely additive under Loewe, but is classified as weak synergy under our metric. The breakpoint experiments, by contrast, did not usually combine antibiotics at precisely half their MIC (since the breakpoint and MIC normalizations differ); as a result, pairs typically needed to exhibit Loewe synergy in order to exhibit synergy according to the ESS. We conclude by recalling that each diagonal was sampled on a twofold dilution gradient, which discretizes the value of Loewe synergies we might observe; if we were to repeat this experiment with a finer discretization, it would increase the precision of the Loewe synergy values.

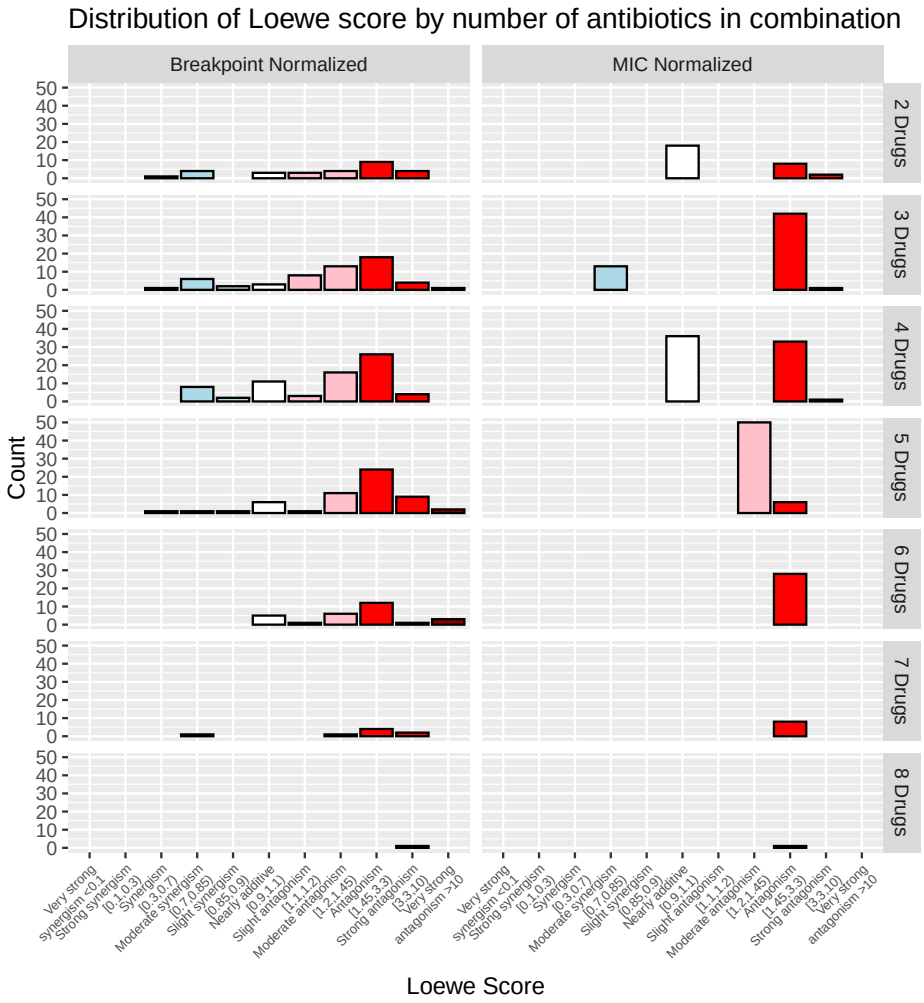


Figure B.2: Loewe interaction scores, calculated along the NDS diagonal, interpreted according to the criteria given in Chou (2006). Observe that the two-fold discretization at which we sampled combinations leads to discretization in realizable Loewe scores, which is most apparent in the MIC-normalized experiment (right panel).

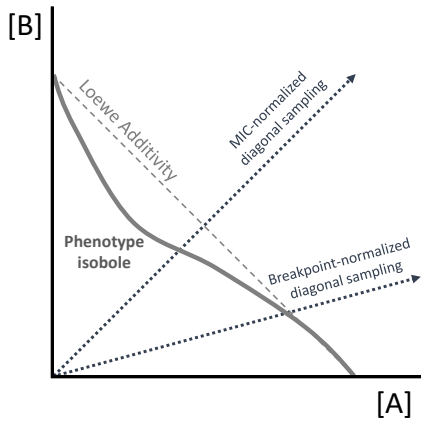


Figure B.3: Depending on the concentration ratio at which Loewe synergy is measured, different degrees of interaction may be reported. In this hypothetical example, sampling at the MIC normalized ratios results in a finding of stronger synergy than the breakpoint normalized ratio.

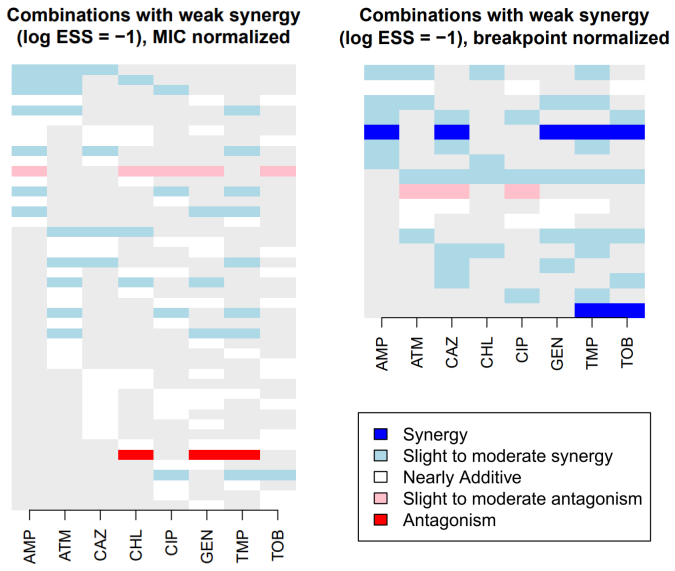


Figure B.4: Loewe synergy determination for the combinations that were determined to have weak synergy ($\log_2 ESS = -1$) according to our metric. Compare to Figures 3.2b and 3.3b; here the "present" antibiotics are colored according to the combination's Loewe interaction.

Appendix C

Appendix for Chapter 4

C.1 Proof of the estimation error lower bound

In this section, we present proofs of lemmas and theorems in Chapter 4.

Setting 1 (Two-covariate Pogit Model)**.** *For $i = 1, 2, \ldots, n$, let covariates $x_{p,i}$ and $x_{\lambda,i}$ be drawn independently according to*

$$x_{\lambda,i} \sim \mathcal{N}(\mu_\lambda, \sigma_\lambda^2) \tag{C.1}$$
$$x_{p,i} \sim \mathcal{N}(0, \sigma_p^2) \tag{C.2}$$

and let Y_i be drawn according to

$$Y_i \sim \textit{Poi}\left(e^{x_{\lambda,i}\theta_\lambda} \frac{\exp\left(x_{p,i}\theta_p\right)}{1 + \exp\left(x_{p,i}\theta_p\right)}\right), \tag{C.3}$$

where $\theta_\lambda, \theta_p \in [C_l, C_u]$ for constants $C_l, C_u \in \mathbb{R}$. The existence of lower and upper bounds on the parameters is needed to prove certain regularity conditions about the maximum likelihood estimator, but the bounds can be chosen such that they are never attained in practical settings.

C.1.1 Regularity conditions

Lemma C.1.1. *Let $\{x_{\lambda,i}, x_{p,i}, Y_i\}_{i=1}^n$ follow* (4.13)*. The following regularity conditions hold.*

1. *θ_0 is identified such that if $\theta \neq \theta_0$ and $\theta \in \Theta$, then $\ell(x, y|\theta) \neq \ell(x, y|\theta_0)$.*
2. *θ_0 lies in the interior of Θ, which is assumed to be a compact subset of $\mathbb{R}^2$.*
3. *$\log \ell(x, y|\theta)$ is cont. differentiable at each $\theta \in \Theta$ a.e. for $x, y \in \mathcal{X} \times \mathcal{Y}$.*
4. *$|\log \ell(x, y|\theta)| \leq d(x, y)$ for all $\theta \in \Theta$ and $\mathbb{E}_{\theta_0}[d(X, Y)] < \infty$.*
5. *$\ell(x, y|\theta)$ is twice cont. differentiable, and $\ell(x, y|\theta) > 0$ in a neighborhood $\mathcal{N}$ of θ_0.*
6. *$||\frac{\partial \ell(x,y|\theta)}{\partial \theta}|| \leq e(x, y)$ for all $\theta \in \mathcal{N}$ and $\int e(x, y) d\nu(x, y) < \infty$.*

7. *Given the score vector* $\psi(x, y|\theta) = (\partial \log \ell(x, y|\theta)/\partial\theta_1, \ldots, \partial \log \ell(x, y|\theta)/\partial\theta_k)'$, *we have that* $\mathcal{I}(\theta_0) = \mathbb{E}_{\theta_0}[\psi(X, Y|\theta_0)\psi(X, Y|\theta_0)']$ *exists and is non-singular.*

8. $||\frac{\partial^2 \log \ell(x,y|\theta)}{\partial\theta\partial\theta'}|| \leq f(x, y)$ *for all* $\theta \in \mathcal{N}$ *and* $\mathbb{E}_{\theta_0}[f(X, Y)] < \infty$.

9. $||\frac{\partial^2 \ell(x,y|\theta)}{\partial\theta\partial\theta'}|| \leq g(x, y)$ *for all* $\theta \in \mathcal{N}$ *and* $\int g(x, y)d\nu(x, y) < \infty$.

And the maximum likelihood estimate of θ *is distributed according to*

$$\hat{\theta}_{MLE} \sim \mathcal{N}\left(\theta, \mathcal{I}(\theta)^{-1}\right). \tag{C.4}$$

We first introduce the likelihood decomposition, which we will use repeatedly in the following proofs:

$$\ell(x, y|\theta) = \ell(y|x, \theta)\ell(x) = \ell(y|\mu(x, \theta))\ell(x) \tag{C.5}$$

And for log likelihood:

$$\log \ell(x, y|\theta) = \log \ell(y|\mu(x, \theta)) + \log \ell(x). \tag{C.6}$$

Condition (1). Identifiability holds because the covariates for p, X_p, are independent from the covariates for λ, X_λ (as shown in Papadopoulos and Santos Silva (2008)).

Condition (2) holds by the assumption given in Setting 1.

Condition (3). From (C.6) and the fact that $\log \ell(y|\mu)$ is continuously differentiable in μ and μ is continuously differentiable in θ, we know that $\log(x, y|\theta)$ is continuously differentiable in θ.

Condition (4). Since Θ is a compact set, for every (x, y) we can attain the maximum and minimum value of $\log \ell(x, y|\theta)$ with respect to θ:

$$\theta_{\max}(x, y) = \max_{\theta\in\Theta} \log \ell(x, y|\theta), \quad \theta_{\min}(x, y) = \min_{\theta\in\Theta} \log \ell(x, y|\theta)$$

And we could write out our upper bound function d in terms of $\theta(x, y)$:

$$d(x, y) = \max\{\log \ell(x, y|\theta)|_{\theta=\theta_{\max}(x,y)}, -\log \ell(x, y|\theta)|_{\theta=\theta_{\min}(x,y)}\}$$

For simplicity, we use $\theta(x, y)$ to ambiguously denote either $\max$ or $\min$ branch.

The expectation in the condition (6) can be written as

$$\begin{aligned}
&\int_x \sum_{y=0}^{\infty} \log \ell(x, y|\theta)|_{\theta=\theta(x,y)} f_Y(y) f_X(x) dx \\
=&\int_x \sum_{y=0}^{\infty} [\log \ell(y|\mu(x, \theta)) + \log \ell(x)]|_{\theta=\theta(x,y)} f_Y(y) f_X(x) dx \\
=&\int_x \log \ell(x) f_X(x) dx + \int_x \sum_{y=0}^{\infty} \log \ell(y|\mu(x, \theta))|_{\theta=\theta(x,y)} f_Y(y) f_X(x) dx
\end{aligned}$$

Here, we argue that since Poisson model satisfies the regularity condition, we know that

$$\sum_{y=0}^{\infty} \log \ell(y|\mu(x,\theta))|_{\theta=\theta(x,y)}\, f_Y(y) \leq \sum_{y=0}^{\infty} \log \ell(y|\mu)|_{\mu=\mu(y)}\, f_Y(y) \leq M_Y.$$

And we assume the distribution of X satisfies the regularity condition, as well:

$$\int_x \log \ell(x) f_X(x) dx \leq M_X.$$

We can therefore conclude that

$$\int_x \sum_{y=0}^{\infty} \log \ell(x,y|\theta)|_{\theta=\theta(x,y)}\, f_Y(y) f_X(x) dx \leq M_X + \int_x M_Y f_X(x) dx = M_X + M_Y.$$

Since the preceding argument applies to both $\theta_{\max}$ and $\theta_{\min}$, we have

$$\mathbb{E}_{\theta_0}[d(X,Y)] < \infty.$$

Condition (5). Using the expression (C.5) and the fact that $\ell(y|\mu)$ is twice differentiable with respect to μ, and μ is twice differentiable with respect to θ, we know that $\ell(x,y|\theta)$ is twice differentiable with respect to θ. Moreover, since $\ell(y|\mu)$ is always positive in the neighborhood of μ and $\ell(x)$ is always positive, we know that $\ell(x,y|\theta)$ is always positive in the neighborhood of θ.

To verify **Condition (6)**, we begin with the decomposition of the likelihood and then use the fact that $\mu(x,\theta) = p(x,\theta)\lambda(x,\theta)$ and (C.5) to write

$$\frac{\partial \ell(x,y|\theta)}{\partial \theta_i} = \ell(x) \frac{\partial \ell(y|\mu(x,\theta))}{\partial \theta_i} = \ell(x) \frac{\partial \ell(y|\mu)}{\partial \mu} \frac{\partial \mu(x,\theta)}{\partial \theta_i}.$$

Since Θ is a compact set, the upper and lower bounds of the partial derivative can be attained. Here, we use same notions $\theta(x,y)$, $\theta_{\max}(x,y)$ and $\theta_{\min}(x,y)$ as in the proof of condition (4). Our dominating function can now be written as

$$e(x,y) = \max \left\{ \ell(x) \frac{\partial \ell(y|\mu)}{\partial \mu} \frac{\partial \mu(x,\theta)}{\partial \theta_i}\bigg|_{\theta=\theta_{\max}(x,y)}, -\ell(x) \frac{\partial \ell(y|\mu)}{\partial \mu} \frac{\partial \mu(x,\theta)}{\partial \theta_i}\bigg|_{\theta=\theta_{\min}(x,y)} \right\}.$$

Since $\nu(x,y)$ is the uniform measure over x and y, we have

$$\int \frac{\partial \ell(x,y|\theta)}{\partial \theta_i}\bigg|_{\theta=\theta(x,y)} d\nu(x,y) = \int_x \ell(x) \frac{\partial \mu(x,\theta)}{\partial \theta_i} \sum_{y=0}^{\infty} \frac{\partial \ell(y|\mu)}{\partial \mu}\bigg|_{\theta=\theta(x,y)} dx.$$

The Poisson likelihood satisfies the regularity condition because it is a generalized linear model, and therefore we have

$$\sum_{y=0}^{\infty} \frac{\partial \ell(y|\mu)}{\partial \mu}\bigg|_{\theta=\theta(x,y)} \leq \sum_{y=0}^{\infty} \frac{\partial \ell(y|\mu)}{\partial \mu}\bigg|_{\mu=\mu(y)} \leq M$$

We can now write an upper bound on the quantity of interest:

$$\int \left.\frac{\partial \ell(x,y|\theta)}{\partial \theta_i}\right|_{\theta=\theta(x,y)} d\nu(x,y) \leq M \int_x \ell(x) \left.\frac{\partial \mu(x,\theta)}{\partial \theta_i}\right|_{\theta=\theta(x,y)} dx.$$

Since $\mu = p \cdot \lambda$ is the product of an exponential and an expit function, the partial derivative of μ with respect to θ can grow at most as fast as exponential function, and it will be dominated by the density function of Gaussian distribution $\ell(x)$. Therefore, we conclude that

$$\int e(x,y) d\nu(x,y) < +\infty,$$

which shows that Condition (6) is satisfied.

Condition (7) is satisfied by inspection of the Fisher Information Matrix, which is computed in the proof of Theorem 4.3.1.

The proof of **Condition (8)** and **Condition (9)** follows from the same arguments used to prove **Condition (4)** and **Condition (6)**, respectively.

Since all conditions are satisfied, the conclusion follows immediately.

C.1.2 Proof of Theorem 4.3.1

We will prove this result using the Cramér-Rao lower bound. The result states that under certain regularity conditions (which we show are satisfied in Lemma C.1.1),

$$\text{Cov}(\hat{\boldsymbol{\theta}}) \succeq \frac{1}{n} \mathcal{I}(\boldsymbol{\theta})^{-1}, , \tag{C.7}$$

where $\mathcal{I}(\theta)$ is the Fisher Information matrix, defined as

$$\mathcal{I}(\boldsymbol{\theta}) := \mathbb{E}_{\boldsymbol{\theta}} \left[\nabla_{\boldsymbol{\theta}} \log \ell_{\boldsymbol{\theta}}(Y,X) \nabla_{\boldsymbol{\theta}} \log \ell_{\boldsymbol{\theta}}(Y,X)^\top \right],$$

where $\ell_{\boldsymbol{\theta}}(Y,X)$ is the likelihood of the observed data under the parameters θ. Under mild regularity conditions, which we show in Lemma C.1.1 are satisfied in this instance, we can write the Fisher information matrix as the negative Hessian of the log likelihood function

$$\mathcal{I}(\boldsymbol{\theta}) = -\mathbb{E}_{\boldsymbol{\theta}} \left[\nabla^2_{\boldsymbol{\theta}} \log \ell_{\boldsymbol{\theta}}(Y,X) \right].$$

We compute the covariance in $\hat{\boldsymbol{\theta}}$ with respect to the randomness in both Y and X:

$$\begin{aligned}
\text{Cov}(\hat{\boldsymbol{\theta}}) &\geq \frac{1}{n} \mathcal{I}^{-1}(\boldsymbol{\theta}) \\
&= \frac{1}{n} \mathbb{E}_{X,Y} \left[-\nabla^2_{\boldsymbol{\theta}} \log \ell_{\boldsymbol{\theta}}(Y,X) \right]^{-1} \\
&= \frac{1}{n} \mathbb{E}_{X,Y} \left[-\nabla^2_{\boldsymbol{\theta}} \left(\log \ell_{\boldsymbol{\theta}}(X) + \log \ell_{\boldsymbol{\theta}}(Y|X) \right) \right]^{-1} \\
&= \frac{1}{n} \mathbb{E}_{X,Y} \left[-\nabla^2_{\boldsymbol{\theta}} \log \ell_{\boldsymbol{\theta}}(Y|X) \right]^{-1},
\end{aligned}$$

where we have used the fact that the distribution of X is independent of the parameters $\boldsymbol{\theta}$. Next, we compute the Hessian of the negative log conditional likelihood. The log conditional likelihood under this model is given by

$$\log \ell_{\boldsymbol{\theta}}(Y|X) = -\log(Y!) - e^{X_\lambda \theta_\lambda}\left(\frac{\exp(X_p\theta_p)}{1+\exp(X_p\theta_p)}\right) + Y\left(X_\lambda\theta_\lambda + \log\left(\frac{\exp(X_p\theta_p)}{1+\exp(X_p\theta_p)}\right)\right),$$

and the Hessian of the negative log conditional likelihood has components

$$-\nabla^2_{\boldsymbol{\theta}} \log \ell_{\boldsymbol{\theta}}(Y|X) = \begin{bmatrix} -\frac{\partial^2}{\partial^2\theta_\lambda}\log\ell_{\boldsymbol{\theta}}(Y|X) & -\frac{\partial^2}{\partial\theta_\lambda\partial\theta_p}\log\ell_{\boldsymbol{\theta}}(Y|X) \\ -\frac{\partial^2}{\partial\theta_p\partial\theta_\lambda}\log\ell_{\boldsymbol{\theta}}(Y|X) & -\frac{\partial^2}{\partial^2\theta_p}\log\ell_{\boldsymbol{\theta}}(Y|X) \end{bmatrix}. \tag{C.8}$$

We will bound each of these quantities independently. Note that the matrix is symmetric; we begin by showing that the off-diagonal entry is zero:

$$\begin{aligned}\mathbb{E}\left[-\frac{\partial^2}{\partial\theta_\lambda\partial\theta_p}\log\ell_{\boldsymbol{\theta}}(Y|X)\right] &= \mathbb{E}\left[e^{X_\lambda\theta_\lambda}\frac{\exp(X_p\theta_p)}{(1+\exp(X_p\theta_p))^2}X_\lambda X_p\right] \\ &= \mathbb{E}\left[e^{X_\lambda\theta_\lambda}X_\lambda\right]\mathbb{E}\left[\frac{\exp(X_p\theta_p)}{(1+\exp(X_p\theta_p))^2}X_p\right].\end{aligned}$$

Note that the second expectation is over an odd function. Since we assumed X_p is symmetric around zero, this term is zero, and

$$\mathbb{E}\left[-\frac{\partial^2}{\partial\theta_\lambda\partial\theta_p}\log\ell_{\boldsymbol{\theta}}(Y|X)\right] = 0.$$

Next, we compute the first diagonal entry in the matrix. We begin by separating it into terms that depend on X_λ and terms that depend on X_p:

$$\begin{aligned}\mathbb{E}\left[-\frac{\partial^2}{\partial^2\theta_\lambda}\log\ell_{\boldsymbol{\theta}}(Y|X)\right] &= \mathbb{E}\left[\sum_i e^{X_\lambda\theta_\lambda}\frac{\exp(X_p\theta_p)}{1+\exp(X_p\theta_p)}X_\lambda^2\right] \\ &= \mathbb{E}\left[e^{X_\lambda\theta_\lambda}X_\lambda^2\right]\mathbb{E}\left[\frac{e^{X_p\theta_p}}{1+e^{X_p\theta_p}}\right] \\ &= \mathbb{E}\left[e^{X_\lambda\theta_\lambda}X_\lambda^2\right]\mathbb{E}[p].\end{aligned}$$

Next, we evaluate the first expectation using the known distribution of X_λ and then completing the square:

$$\begin{aligned}\mathbb{E}\left[-\frac{\partial^2}{\partial^2\theta_\lambda}\log\ell_{\boldsymbol{\theta}}(Y|X)\right] &= \mathbb{E}[p]\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma_\lambda^2}}x^2\exp\left(-\frac{(x-\mu_\lambda)^2}{2\sigma_\lambda^2}+x\theta_\lambda\right)dx \\ &= \mathbb{E}[p]\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi\sigma_\lambda^2}}x^2 e^{\mu_\lambda\theta_\lambda+\sigma_\lambda^2\theta_\lambda^2/2}\exp\left(-\frac{1}{2\sigma_\lambda^2}(x-(\mu_\lambda+\sigma_\lambda^2\theta_\lambda))^2\right)dx \\ &= \mathbb{E}[p]e^{\mu\theta_\lambda+\sigma_\lambda^2\theta_\lambda^2/2}\mathbb{E}_{x\sim\mathcal{N}(\mu_\lambda+\sigma_\lambda^2\theta_\lambda,\sigma_\lambda^2)}[x^2] \\ &= \mathbb{E}[p]\mathbb{E}[\lambda]\left((\mu_\lambda+\sigma_\lambda^2\theta_\lambda)^2+\sigma_\lambda^2\right).\end{aligned}$$

Now, all that remains is to upper bound the final term in the Hessian of the negative log likelihood:

$$\mathbb{E}\left[-\frac{\partial^2}{\partial^2\theta_p}\log \ell_{\boldsymbol{\theta}}(Y|X)\right] = \mathbb{E}\left[\frac{\exp(X_p\theta_p)}{(1+\exp(X_p\theta_p))^3}\left(e^{X_\lambda\theta_\lambda}\left(1-e^{X_p\theta_p}\right)+Y\left(1+e^{X_p\theta_p}\right)\right)X_pX_p\right].$$

We begin with the tower rule of expectation, using the fact that $\mathbb{E}[Y|X] = \frac{\exp(X_p\theta_p)}{1+\exp(X_p\theta_p)}e^{X_\lambda\theta_\lambda}$.

$$\begin{aligned}
\mathbb{E}\left[-\frac{\partial^2}{\partial^2\theta_p}\log \ell_{\boldsymbol{\theta}}(Y|X)\right] &= \mathbb{E}\left[\mathbb{E}\left[\frac{\exp(X_p\theta_p)}{(1+\exp(X_p\theta_p))^3}\left(e^{X_\lambda\theta_\lambda}\left(1-e^{X_p\theta_p}\right)+Y\left(1+e^{X_p\theta_p}\right)\right)X_pX_p\middle|X\right]\right] \\
&= \mathbb{E}\left[\frac{\exp(X_p\theta_p)}{(1+\exp(X_p\theta_p))^3}\left(e^{X_\lambda\theta_\lambda}\left(1-e^{X_p\theta_p}\right)+\mathbb{E}\left[Y|X\right]\left(1+e^{X_p\theta_p}\right)\right)X_p^2\right] \\
&= \mathbb{E}\left[e^{X_\lambda\theta_\lambda}\frac{\exp(X_p\theta_p)}{(1+\exp(X_p\theta_p))^3}X_p^2\right]
\end{aligned}$$

Next, we use the fact that X_p and X_λ are independent to separate the expectation over X_p from the expectation over X_λ:

$$\begin{aligned}
\mathbb{E}\left[-\frac{\partial^2}{\partial^2\theta_p}\log \ell_{\boldsymbol{\theta}}(Y|X)\right] &= \mathbb{E}\left[e^{X_\lambda\theta_\lambda}\right]\mathbb{E}\left[\frac{\exp(X_p\theta_p)}{(1+\exp(X_p\theta_p))^3}X_p^2\right] \\
&= \mathbb{E}[\lambda]\mathbb{E}\left[\frac{\exp(X_p\theta_p)}{(1+\exp(X_p\theta_p))^3}X_p^2\right].
\end{aligned}$$

We will now show that, *regardless of the distribution of* X_p, this remaining expectation is less than $\frac{1}{2\theta_p^2}$. Our strategy will be to upper bound the expectation by the maximum value of its argument. We have

$$\begin{aligned}
\mathbb{E}\left[\frac{\exp(X_p\theta_p)}{(1+\exp(X_p\theta_p))^3}X_p^2\right] &\leq \max_x \frac{\exp(x\theta_p)}{(1+\exp(x\theta_p))^3}x^2 \\
&= \frac{1}{\theta_p^2}\max_x \frac{\exp(x\theta_p)}{(1+\exp(x\theta_p))^3}(x\theta_p)^2.
\end{aligned}$$

Now, let $u = x\theta_p$:

$$\mathbb{E}\left[\frac{\exp(X_p\theta_p)}{(1+\exp(X_p\theta_p))^3}X_p^2\right] \leq \frac{1}{\theta_p^2}\max_u \frac{\exp(u)}{(1+\exp(u))^3}u^2.$$

Here, we provide a simple upper bound for quantity

$$C \triangleq \max_u \frac{\exp(u)}{(1+\exp(u))^3}u^2.$$

We divide the maximization problem over cases when u is positive and when u is non-positive. When $u > 0$, we know that

$$\frac{\exp(u)}{(1+\exp(u))^3}u^2 = \frac{\exp(u)}{(1+\exp(u))^2}\frac{u^2}{1+\exp(u)} \leq \frac{\exp(u)}{(1+\exp(u))^2} \leq \frac{1}{4}.$$

And when $u \leq 0$, we have

$$\begin{aligned}\frac{\exp(u)}{(1+\exp(u))^3}u^2 &= \frac{1}{1+\exp(u)}\frac{u^2}{(1+\exp(u))(1+\exp(-u))} \\ &= \frac{1}{1+\exp(u)}\frac{u^2}{2+\exp(u)+\exp(-u)} \\ &\leq \frac{u^2}{2+\exp(u)+\exp(-u)} \\ &\leq \frac{u^2}{4+u^2+\frac{u^4}{12}} \leq \frac{\sqrt{3}}{\sqrt{3}+2} \leq \frac{1}{2}.\end{aligned}$$

Therefore, we know that $C \leq 1/2$.

We conclude that, regardless of the distribution of X_p, we have

$$\mathbb{E}\left[-\frac{\partial^2}{\partial^2\theta_p}\log \ell_{\boldsymbol{\theta}}(Y|X)\right] \leq \tfrac{1}{2}\mathbb{E}[\lambda]\theta_p^{-2}.$$

We have shown that the Fisher information matrix is diagonal and lower bounded its diagonal entries, so the conclusion of the theorem follows from Eqn (C.7).

Appendix D

Appendix for Chapter 5

D.1 High variance of the IPS estimator

In this section we substantiate the claim from Section 5.2.3 that the inverse propensity score (IPS) estimator is too high-variance to be practical in the settings we consider. The core idea is that, when we run one-sided bipartite experiments, the definition of "complete exposure" to treatment or control is defined based on a unit's two-hop neighborhood. In a graph with even moderate connectivity, these neighborhoods are so large that units have a very low chance of being fully exposed to treatment or control. These small probabilities of full exposure in turn lead to high variance of the IPS estimator.

We begin by observing that, using cluster-randomized designs as we have defined them in 5.3.1, there is essentially zero chance of each unit being connected completely to treatment or control under even moderate connectivity. To see this, we will compute the chance of a given experimental unit i being a two-hop neighbor of unit j under the bipartite stochastic block model discussed in Section 5.5.1. Recall that in this model, N experimental units and M interference units are each partitioned into K equal-sized clusters labeled $1, \ldots, K$. Units with the same label have an edge between them with probability $q = 0.5$, while units with different labels have an edge with probability p. We can compute the chance of experimental unit i **not** having a two-hop neighbor in $\mathcal{C}(j)$ if $\mathcal{C}(i) \neq \mathcal{C}(j)$ is given by

$$\begin{aligned}
&\mathbb{P}(i \text{ has no 2-hop neighbor in } \mathcal{C}(j) | \mathcal{C}(i) \neq \mathcal{C}(j)) = \\
&\qquad \prod_{k \in \mathcal{C}(j)} \prod_{s \in \mathcal{C}(j)} \mathbb{P}(k \text{ not connected to } s \text{ OR } i \text{ not connected to } s) \cdot \\
&\qquad\qquad \prod_{k \in \mathcal{C}(j)} \prod_{s \in \mathcal{C}(i)} \mathbb{P}(k \text{ not connected to } s \text{ OR } i \text{ not connected to } s) \\
&\qquad = \prod_{k \in \mathcal{C}(j)} \prod_{s \in \mathcal{C}(j)} (1 - pq) \cdot \prod_{k \in \mathcal{C}(j)} \prod_{s \in \mathcal{C}(i)} (1 - pq) \\
&\qquad = (1 - pq)^{2MN/K^2}
\end{aligned}$$

The number of clusters to which unit i is connected, besides $\mathcal{C}(i)$, is the binomial random variable $\text{Binom}\left(K - 1, 1 - (1 - pq)^{2MN/K^2}\right)$. Observe that the probability of connection grows very quickly. In the setting of our experiments in Section 5.5.1, we have $M = 2000$, $N = 1000$, $q = 0.5$ and $K = 20$, so that when $p = 0.005$ we have $\mathbb{P}(i \text{ has a two-hop neighbor in } \mathcal{C}(j) | \mathcal{C}(i) \neq \mathcal{C}(j)) =$

Table D.1: RMSE (relative to τ) of $\widehat{\tau}_{DIM}$ and $\widehat{\tau}_{IPS}$ as the neighborhood of pure exposure, Δ, widens (see 5.5.2)

	$\Delta = 0.1$	$\Delta = 0.3$	$\Delta = 0.5$
$\widehat{\tau}_{DIM}$	1.001	0.459	0.001
$\widehat{\tau}_{IPS}$	0.429	0.043	0.032

$1 - 1 \cdot 10^{-11}$, i.e. each unit is almost certain to have a two-hop neighbor in every other cluster. We conclude that, in the graph settings we study, experimental units are likely to be connected to units of all clusters, making pure exposure to treatment or control an exceedingly rare occurrence.

To further illustrate this point, we compute the variance of the IPS estimator on our simulated graphs for various definitions of "pure exposure". Under Bernoulli cluster randomization (which is slightly different from the design studied in the rest of Chapter 5, but which is standard for the IPS estimator) we have:

$$\text{Var}(\widehat{\tau}_{IPS}) = \frac{1}{N^2} \sum_{i \in [N]} \frac{1}{\mathbb{P}(\text{unit } i \text{ is treated})} Y_{i,T}^2 + \frac{1}{\mathbb{P}(\text{unit } i \text{ is controlled})} Y_{i,C}^2$$

where $Y_{i,T}$ is the value of Y_i under full treatment, and $Y_{i,C}$ is the value of Y_i under full control. We compute the probabilities of full treatment and control via Monte Carlo simulation over draws of $\mathbf{Z}$ from a Bernoulli randomized design.

We compute the variance of the IPS estimator in two settings studied in Section 5.5.

Table D.1 compares the bias of $\widehat{\tau}_{DIM}$ to the standard deviation of $\widehat{\tau}_{IPS}$ in the setting of Section 5.5.2, in which units behave as if they were fully exposed whenever $|Z_i - e_i| < \Delta$. This is precisely the setting in which the $\widehat{\tau}_{IPS}$ estimator works best, and it is unbiased in this case. We see from the table that the IPS estimator has less error than $\widehat{\tau}_{DIM}$ for small values of Δ, but that when $\Delta = 0.5$, the bias of $\widehat{\tau}_{DIM}$ has decreased to be lower than the IPS variance. We note that these figures were generated for the graph with $p = 0.005$; increasing the connectivity (say, to $p = 0.05$) further increases the error of $\widehat{\tau}_{IPS}$ relative to $\widehat{\tau}_{DIM}$.

Table D.2 compares $\widehat{\tau}_{IPS}$ and $\widehat{\tau}_{DIM}$ in the linear setting of Section 5.5.1. In this setting, the IPS estimator relies on the incorrect assumption that units in a Δ neighborhood of pure exposure act as if they were purely exposed to treatment or control; this is untrue in the linear model, where even slight exposure to the opposite treatment results in a change in Y_i. As a result, the standard deviation reported in Table D.2 provides a lower bound on the RMSE of $\widehat{\tau}_{IPS}$, with the remainder of the error due to bias. We see that the IPS estimator with $\Delta = 0.1$ and $\Delta = 0.3$ has relative error that is much higher than $\widehat{\tau}_{DIM}$, even before including the bias of $\widehat{\tau}_{IPS}$. When $\Delta = 0.5$, we expect the bias of $\widehat{\tau}_{IPS}$ to be substantial.

We conclude that, even though $\widehat{\tau}_{IPS}$ provides an unbiased estimate of τ when the exposure mapping is correctly specified, the variance in this estimator can be significant enough to justify

Table D.2: Bias (relative to τ) of $\widehat{\tau}_{DIM}$ and standard deviation (relative to τ) of $\widehat{\tau}_{IPS}$ as the bipartite stochastic block model changes (see 5.5.1)

	$p = 0.0$	$p = 0.005$	$p = 0.05$	$p = 0.5$
$\widehat{\tau}_{DIM}$	0.01	3.88	11.58	12.96
$\widehat{\tau}_{IPS}(\Delta = 0.1)$	0.70	9.88	390.63	508.78
$\widehat{\tau}_{IPS}(\Delta = 0.3)$	0.70	9.5	13.90	16.54
$\widehat{\tau}_{IPS}(\Delta = 0.5)$	0.70	0.71	3.70	4.31

using the biased estimator $\widehat{\tau}_{DIM}$.

D.2 Proofs

In this appendix, we provide proofs of all the results in Chapter 5. We apply similar proof techniques to bound the bias in a variety of potential outcomes models, and therefore consider a potential outcomes model which generalizes all the models discussed in Chapter 5:

$$Y_i = g_i(Z_i, e_i).$$

We will first prove some useful lemmas that apply to all potential outcomes models of this form; later we will provide results that are specific to each model type.

D.2.1 Useful Lemmas

The first two lemmas in this section will be useful for calculating the bias of the difference in means estimator under the linear, Lipschitz, and $\Delta-$neighborhood functions. The proofs for each bias bound will follow a similar structure: First, we will decompose the bias of the difference-in-means estimator into the bias contribution of each unit under both treatment and control. Next, we will use the given structure of the potential outcome to either compute the bias exactly (in the linear setting) or bound the bias (in the Lipschitz and $\Delta-$neighborhood settings). In each case, we will be left with a bound in terms of the conditional expectation of the exposure e_i given the treatment status Z_i. Our third step will be to relate this quantity to the folded graph clustering objective, $\mathcal{H}(\mathcal{C})$.

We provide lemmas for the first and third steps; the second step is unique to each bias calculation.

Lemma D.2.1 (Unit-level bias decomposition)**.** *Let the unit-level responses Y_i be functions of the unit's treatment Z_i and the unit-level exposure e_i, so that $Y_i = g_i(Z_i, e_i)$. Then the bias of the*

difference-in-means estimate of the average treatment effect can be written as

$$\mathbb{E}[\hat{\tau}] - \tau^* = \frac{1}{N}\sum_{i\in[N]} \mathbb{E}[g_i(Z_i, e_i) - g_i(Z_i, Z_i)|Z_i = 1] - \frac{1}{N}\sum_{i\in[N]} \mathbb{E}[g_i(Z_i, e_i) - g_i(Z_i, Z_i)|Z_i = -1]$$

Proof. We begin by writing the bias in terms of the response function $g_i(Z_i, e_i)$:

$$\begin{aligned}
\mathbb{E}[\hat{\tau}] - \tau^* &= \mathbb{E}\left[\frac{1}{N_T}\sum_{i\in I} Y_i - \frac{1}{N_C}\sum_{i\in \bar{I}} Y_i\right] - \frac{1}{N}\sum_{i\in[N]}(g_i(1,1) - g_i(-1,-1)) \\
&= \frac{1}{N_T}\sum_{i\in[N]} \mathbb{E}\left[\mathbf{1}\{Z_i = 1\}Y_i\right] - \frac{1}{N_C}\sum_{i\in[N]} \mathbb{E}\left[\mathbf{1}\{Z_i = -1\}Y_i\right] - \frac{1}{N}\sum_{i\in[N]}(g_i(1,1) - g_i(-1,-1)) \\
&= \frac{1}{N_T}\sum_{i\in[N]} \mathbb{E}\left[\mathbf{1}\{Z_i = 1\}g_i(Z_i, e_i)\right] - \frac{1}{N_C}\sum_{i\in[N]} \mathbb{E}\left[\mathbf{1}\{Z_i = -1\}g_i(Z_i, e_i)\right] - \\
&\qquad \frac{1}{N}\sum_{i\in[N]}(g_i(1,1) - g_i(-1,-1))
\end{aligned}$$

We use the law of total probability to rewrite the expectations as conditional expectations:

$$\begin{aligned}
\mathbb{E}[\hat{\tau}] - \tau^* &= \frac{1}{N_T}\sum_{i\in[N]} \mathbb{E}\left[g_i(Z_i, e_i)|Z_i = 1\right]\mathbb{P}(Z_i = 1) - \frac{1}{N_C}\sum_{i\in[N]} \mathbb{E}\left[g_i(Z_i, e_i)|Z_i = -1\right]\mathbb{P}(Z_i = -1) \\
&\qquad - \frac{1}{N}\sum_{i\in[N]}(g_i(1,1) - g_i(-1,-1)) \\
&= \frac{1}{N_T}\sum_{i\in[N]} \frac{N_T}{N}\mathbb{E}\left[g_i(Z_i, e_i)|Z_i = 1\right] - \frac{1}{N_C}\sum_{i\in[N]} \frac{N_C}{N}\mathbb{E}\left[g_i(Z_i, e_i)|Z_i = -1\right] - \\
&\qquad \frac{1}{N}\sum_{i\in[N]}(g_i(1,1) - g_i(-1,-1)) \\
&= \frac{1}{N}\sum_{i\in[N]} \mathbb{E}\left[g_i(Z_i, e_i)|Z_i = 1\right] - \frac{1}{N}\sum_{i\in[N]} \mathbb{E}\left[g_i(Z_i, e_i)|Z_i = -1\right] - \\
&\qquad \frac{1}{N}\sum_{i\in[N]}(g_i(1,1) - g_i(-1,-1))
\end{aligned}$$

All of the summations are now computing averages over the N terms. We can distribute the terms of the final summation between the first and second summations, which lets us decompose the

bias into contributions from each unit under its control and treated assignments:

$$\begin{aligned}\mathbb{E}[\hat{\tau}] - \tau^* &= \frac{1}{N}\sum_{i\in[N]} \mathbb{E}\left[g_i(Z_i, e_i) - g_i(1,1)|Z_i = 1\right] - \frac{1}{N}\sum_{i\in[N]} \mathbb{E}\left[g_i(Z_i, e_i) - g_i(-1,-1)|Z_i = -1\right] \\ &= \frac{1}{N}\sum_{i\in[N]} \mathbb{E}\left[g_i(Z_i, e_i) - g_i(Z_i, Z_i)|Z_i = 1\right] - \frac{1}{N}\sum_{i\in[N]} \mathbb{E}\left[g_i(Z_i, e_i) - g_i(Z_i, Z_i)|Z_i = -1\right]\end{aligned}$$

as desired. □

Lemma D.2.2 (Writing the weighted condition gaps in terms of the graph structure)**.** *Let e_i be the exposure of unit i to treatment, as defined in Eqn* (5.2)*. Let $\gamma_i \in \mathbb{R}$ be arbitrary. If the treatment assignment vector $\mathbf{Z} \sim \mathcal{D}(\mathcal{C})$ is drawn according to a balanced cluster randomized design (definition 5.3.1), then the γ_i-weighted average conditional gap between e_i and the unit's exposure Z_i can be written in terms of the underlying graph weights w_{is} between units assigned to different clusters:*

$$\frac{1}{N}\sum_{i\in[N]} \gamma_i \left(\mathbb{E}\left[Z_i - e_i|Z_i = 1\right] + \mathbb{E}\left[e_i - Z_i|Z_i = -1\right]\right) = \frac{2}{N}\frac{K}{K-1}\sum_{i\in[N]}\sum_{j\notin\mathcal{C}(i)} \gamma_i \sum_s \frac{w_{is}}{\sum_s w_{is}}\frac{w_{js}}{\sum_k w_{ks}}.$$

Proof. We begin by observing the useful fact that

$$\sum_s \frac{w_{is}}{\sum_s w_{is}} \sum_j \frac{w_{js}}{\sum_j w_{js}} = 1.$$

Combining this fact with the definition of e_i lets us write all of the terms in this expression as linear combinations of Z_j and Z_i, where Z_i is fixed in each conditional expectation.

$$\begin{aligned}&\frac{1}{N}\sum_{i\in[N]} \gamma_i \left(\mathbb{E}\left[Z_i - e_i|Z_i = 1\right] + \mathbb{E}\left[e_i - Z_i|Z_i = -1\right]\right) \\ &\quad= \frac{1}{N}\sum_{i\in[N]} \gamma_i \left(\mathbb{E}\left[\sum_s \frac{w_{is}}{\sum_s w_{is}} \sum_j \frac{w_{js}}{\sum_j w_{js}}(Z_i - Z_j)|Z_i = 1\right]\right. \\ &\qquad\qquad \left. + \mathbb{E}\left[\sum_s \frac{w_{is}}{\sum_s w_{is}} \sum_j \frac{w_{js}}{\sum_j w_{js}}(Z_j - Z_i)|Z_i = -1\right]\right) \\ &\quad= \frac{1}{N}\sum_{i\in[N]} \gamma_i \sum_s \frac{w_{is}}{\sum_s w_{is}} \sum_j \frac{w_{js}}{\sum_j w_{js}} \left(\mathbb{E}\left[(Z_i - Z_j)|Z_i = 1\right] + \mathbb{E}\left[(Z_j - Z_i)|Z_i = -1\right]\right) \\ &\quad= \frac{1}{N}\sum_{i\in[N]} \gamma_i \sum_s \frac{w_{is}}{\sum_s w_{is}} \sum_j \frac{w_{js}}{\sum_j w_{js}} \left(2\mathbb{P}\left(Z_j = -1|Z_i = 1\right) + 2\mathbb{P}\left(Z_j = 1|Z_i = -1\right)\right)\end{aligned}$$

where in the last step we used the fact that Z_j only takes values in $\{-1, 1\}$.

Recall that the Z_i were assigned according to a *cluster-randomized* design $\mathcal{C}$. If $j \in \mathcal{C}(i)$ then $Z_i = Z_j$, so that units in the same cluster contribute zero to the summation above. Otherwise, if

$j \notin \mathcal{C}(i)$, we can compute the probability that $Z_i \neq Z_j$. If there are K clusters and K_T of them are chosen to be treated, then we have

$$\mathbb{P}(Z_i = -1 | Z_j = 1 \cap j \notin \mathcal{C}(i)) = \frac{K_C}{K-1}$$

and

$$\mathbb{P}(Z_i = 1 | Z_j = -1 \cap j \notin \mathcal{C}(i)) = \frac{K_T}{K-1}.$$

We can use this information to compute the probabilities in the expression above, applying the fact that $\mathbb{P}(Z_i \neq Z_j | j \in \mathcal{C}(i)) = 0$ to restrict the sum over j to only the units outside of $\mathcal{C}(i)$.

$$\begin{aligned}
&\frac{1}{N} \sum_{i \in [N]} \gamma_i \left(\mathbb{E}\left[Z_i - e_i | Z_i = 1 \right] + \mathbb{E}\left[e_i - Z_i | Z_i = -1 \right] \right) \\
&\quad = \frac{2}{N} \sum_{i \in [N]} \gamma_i \sum_s \frac{w_{is}}{\sum_s w_{is}} \sum_{j \notin \mathcal{C}(i)} \frac{w_{js}}{\sum_j w_{js}} \left(\mathbb{P}\left(Z_j = -1 | Z_i = 1 \cap j \notin \mathcal{C}(i) \right) + \mathbb{P}\left(Z_j = 1 | Z_i = -1 \cap j \notin \mathcal{C}(i) \right) \right) \\
&\quad = \frac{2}{N} \sum_{i \in [N]} \gamma_i \sum_s \frac{w_{is}}{\sum_s w_{is}} \sum_{j \notin \mathcal{C}(i)} \frac{w_{js}}{\sum_j w_{js}} \left(\frac{K_C}{K-1} + \frac{K_T}{K-1} \right) \\
&\quad = \frac{2}{N} \sum_{i \in [N]} \gamma_i \sum_s \frac{w_{is}}{\sum_s w_{is}} \sum_{j \notin \mathcal{C}(i)} \frac{w_{js}}{\sum_j w_{js}} \frac{K}{K-1},
\end{aligned}$$

as desired. □

The final lemma in this section writes the exposure vector $\mathbf{e}$ is a linear combination of the treatment assignments $\mathbf{Z}$. This linearity is useful for establishing the connection between the covariance objective and the folded graph objective in Lemma 5.3.3.

Lemma D.2.3. *Let* $\mathbf{e}$ *be defined as in Equation* (5.2)*. Then* $\mathbf{e} = C\mathbf{Z}$*, where*

$$C_{ij} = \sum_s \frac{w_{is}}{\sum_s w_{is}} \frac{w_{js}}{\sum_k w_{ks}}$$

Proof. The proof proceeds by the definition of e_i. Let matrix $B \in [0,1]^{N \times M}$ be defined as

$$B_{is} = \frac{w_{is}}{\sum_s w_{is}}.$$

Then we can write the exposures e_i as a linear combination of the doses d_s:

$$\begin{aligned}
e_i &= \frac{\sum_s w_{is} d_s}{\sum_s w_{is}} \\
&= [B\mathbf{d}]_i.
\end{aligned}$$

Similarly, let matrix $A \in [0,1]^{M \times N}$ be defined as

$$A_{si} = \frac{w_{is}}{\sum_i w_{is}},$$

so that we can write the doses d_s as a linear combination of the treatment effects $\mathbf{Z}$:

$$\begin{aligned} d_s &= \frac{\sum_i w_{is} Z_i}{\sum_i w_{is}} \\ &= [A\mathbf{Z}]_s. \end{aligned}$$

Putting these together, we have

$$\begin{aligned} \mathbf{e} &= BA\mathbf{Z} \\ &=: C\mathbf{Z} \end{aligned}$$

with C_{ij} as given in the lemma statement. □

With these helper lemmas established, we turn to proving the main results in Chapter 5.

D.2.2 Proof of Lemma 5.3.2

We begin by applying Lemma D.2.1 to decompose the bias in terms of its unit-level contributions:

$$\mathbb{E}[\hat{\tau}] - \tau^* = \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[g_i(Z_i, e_i) - g(Z_i, Z_i) | Z_i = 1\right] - \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[g_i(Z_i, e_i) - g(Z_i, Z_i) | Z_i = -1\right] \tag{D.1}$$

We apply the linear response function to simplify this expression:

$$\begin{aligned} \mathbb{E}[\hat{\tau}] - \tau^* &= \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[(\alpha_i + \beta_i Z_i + \gamma_i e_i) - (\alpha_i + \beta_i Z_i + \gamma_i Z_i) | Z_i = 1\right] && \text{(D.2)} \\ &\quad - \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[(\alpha_i + \beta_i Z_i + \gamma_i e_i) - (\alpha_i + \beta_i Z_i + \gamma_i Z_i) | Z_i = -1\right] && \text{(D.3)} \\ &= \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[\gamma_i e_i - \gamma_i Z_i | Z_i = 1\right] - \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[\gamma_i e_i - \gamma_i Z_i | Z_i = -1\right] && \text{(D.4)} \\ &= -\frac{1}{N} \sum_{i \in [N]} \gamma_i \left(\mathbb{E}\left[Z_i - e_i | Z_i = 1\right] + \mathbb{E}\left[e_i - (-1) | Z_i = -1\right]\right) && \text{(D.5)} \end{aligned}$$

Next, we apply Lemma D.2.2 to write this expression in terms of the graph structure, completing the proof of the first statement of the lemma:

$$\mathbb{E}[\widehat{\tau}] - \tau^* = -\frac{2}{N} \frac{K}{K-1} \sum_{i \in [N]} \sum_{j \notin \mathcal{C}(i)} \gamma_i \sum_s \frac{w_{is}}{\sum_s w_{is}} \frac{w_{js}}{\sum_k w_{ks}}. \tag{D.6}$$

Next, we will prove that the folded graph objective provides the bias-minimizing clustering in a minimax sense among all balanced cluster-randomized designs when the potential outcome is linear in Z and the exposure e, and the interference parameter γ is bounded on a shared interval for all i.

We begin by finding the maximum bias (over choice of γ) for a given clustering $\mathcal{C}$. We use the value of the bias from Eqn (D.6):

$$\arg\min_{\mathcal{C}} \max_{\gamma \in [\Gamma_0, \Gamma_1]} \left| \mathbb{E}_{\mathbf{Z} \sim \mathcal{D}(\mathcal{C})}[\hat{\tau}] - \tau^* \right| = \arg\min_{\mathcal{C}} \max_{\gamma \in [\Gamma_0, \Gamma_1]^N} 2\frac{1}{N} \cdot \frac{K}{K-1} \left| \sum_{i \in [N]} \sum_{j \notin \mathcal{C}(i)} \gamma_i \sum_s \frac{w_{is}}{\sum_s w_{is}} \frac{w_{js}}{\sum_k w_{ks}} \right|. \tag{D.7}$$

Recall that the edge weights w_{is} are all nonnegative, so term i of the summation takes on the sign of γ_i. For this reason, the maximum bias occurs when all γ_i are of the same sign (so that no terms in the summation cancel each other), at $\gamma_i = \max(|\Gamma_0|, |\Gamma_1|)$.

$$\arg\min_{\mathcal{C}} \max_{\gamma \in [\Gamma_0, \Gamma_1]} \left| \mathbb{E}_{\mathbf{Z} \sim \mathcal{D}(\mathcal{C})}[\hat{\tau}] - \tau^* \right| = \arg\min_{\mathcal{C}} \max(|\Gamma_0|, |\Gamma_1|) \cdot 2\frac{1}{N} \cdot \frac{K}{K-1} \sum_{i \in [N]} \sum_{j \notin \mathcal{C}(i)} \sum_s \frac{w_{is}}{\sum_s w_{is}} \frac{w_{js}}{\sum_k w_{ks}}. \tag{D.8}$$

The terms K, N, Γ_0 and Γ_1 are all constants with respect to the clustering $\mathcal{C}$, so they can be removed without affecting the arg max. We recognize this objective as precisely our folded graph objective

$$\arg\min_{\mathcal{C}} \max_{\gamma \in [\Gamma_0, \Gamma_1]} \left| \mathbb{E}_{\mathbf{Z} \sim \mathcal{D}(\mathcal{C})}[\hat{\tau}] - \tau^* \right| = \arg\min_{\mathcal{C}} H(\mathcal{C}) \tag{D.9}$$

as desired. □

D.2.3 Proof of Lemma 5.4.1 (bounding the bias under the Lipschitz potential outcomes model)

We begin by applying Lemma D.2.1 to decompose the bias in terms of its unit-level contributions:

$$\begin{aligned} |\mathbb{E}[\hat{\tau}] - \tau^*| &= \left| \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[g_i(Z_i, e_i) - g(Z_i, Z_i) | Z_i = 1 \right] - \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[g_i(Z_i, e_i) - g(Z_i, Z_i) | Z_i = -1 \right] \right| \\ &= \left| \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[g_i(1, e_i) - g(1, Z_i) | Z_i = 1 \right] - \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[g_i(-1, e_i) - g(-1, Z_i) | Z_i = -1 \right] \right| \end{aligned}$$

Next, we use the fact that $g_i(Z, e)$ is L-Lipschitz in e to simplify this expression:

$$\begin{aligned} |\mathbb{E}[\hat{\tau}] - \tau^*| &\leq \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[L \left| Z_i - e_i \right| \middle| Z_i = 1 \right] + \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[L \left| e_i - Z_i \right| \middle| Z_i = -1 \right] \\ &= \frac{1}{N} \sum_{i \in [N]} L\mathbb{E}\left[\left| 1 - e_i \right| \middle| Z_i = 1 \right] + \frac{1}{N} \sum_{i \in [N]} L\mathbb{E}\left[\left| e_i - (-1) \right| \middle| Z_i = -1 \right]. \end{aligned}$$

Observe that $e_i \in [-1, 1]$, so that the terms in absolute values are all positive. This lets us drop the absolute value signs and write

$$|\mathbb{E}[\hat{\tau}] - \tau^*| \leq \frac{1}{N} \sum_{i \in [N]} L\mathbb{E}\left[1 - e_i | Z_i = 1\right] + \frac{1}{N} \sum_{i \in [N]} L\mathbb{E}\left[e_i - (-1) | Z_i = -1\right].$$

We apply Lemma D.2.2 with $\gamma_i = L$ to bound the bias as

$$|\mathbb{E}[\hat{\tau}] - \tau^*| \leq \frac{2}{N} \frac{K}{K-1} L \sum_{i \in [N]} \sum_{j \notin \mathcal{C}(i)} \frac{1}{\sum_s w_{is}} w_i^T \tilde{w}_j.$$

To prove the second part of the lemma statement (minimax optimality), we compute the maximum bias over L-Lipschitz interference functions g_i, using the expression for bias given in Equation (D.10).

$$\begin{aligned} &\arg\min_{\mathcal{C}} \max_{\{g_i \in \text{Lip}_L(e)\}} |\mathbb{E}[\hat{\tau}] - \tau^*| \qquad \text{(D.10)} \\ &\quad = \arg\min_{\mathcal{C}} \max_{\{g_i \in \text{Lip}_L(e)\}} \Bigg| \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[g_i(1,1) - g_i(1, e_i) | Z_i = 1\right] + \\ &\qquad\qquad \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[g_i(-1, e_i) - g_i(-1, -1) | Z_i = -1\right] \Bigg|. \end{aligned}$$

Next, we will show that setting $g_i(Z, e) = L \cdot e \ \forall i$ achieves the maximum over all L-Lipschitz functions by computing an upper bound on the argument of the maximum and showing that this choice of $\{g_i\}$ attains that bound. We begin by upper bounding the maximum:

$$\begin{aligned} &\max_{\{g_i \in \text{Lip}_L(e)\}} \left| \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[g_i(1,1) - g_i(1, e_i) | Z_i = 1\right] + \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[g_i(-1, e_i) - g_i(-1, -1) | Z_i = -1\right] \right| \\ &\quad \leq \max_{\{g_i \in \text{Lip}_L(e)\}} \Bigg(\frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[\left| g_i(1,1) - g_i(1, e_i) \right| \Big| Z_i = 1 \right] + \\ &\qquad\qquad \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[\left| g_i(-1, e_i) - g_i(-1, -1) \right| \Big| Z_i = -1 \right] \Bigg) \\ &\quad \leq \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[\max_{g_i \in \text{Lip}_L(e)} \left| g_i(1,1) - g_i(1, e_i) \right| \Big| Z_i = 1 \right] + \\ &\qquad \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[\max_{g_i \in \text{Lip}_L(e)} \left| g_i(-1, e_i) - g_i(-1, -1) \right| \Big| Z_i = -1 \right] \end{aligned}$$

Next we apply the Lipschitz assumption, the fact that $e_i \in [-1, 1]$, and Lemma D.2.2 with $\gamma_i = L$:

$$\begin{aligned}
&\max_{\{g_i \in \text{Lip}_L(e)\}} \left| \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[g_i(1,1) - g_i(1,e_i) | Z_i = 1\right] + \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[g_i(-1,e_i) - g_i(-1,-1) | Z_i = -1\right] \right| \\
&\quad \leq \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[L\left|1 - e_i\right| \Big| Z_i = 1\right] + \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[L\left|e_i - (-1)\right| \Big| Z_i = -1\right] \\
&\quad = \frac{1}{N} \sum_{i \in [N]} L\mathbb{E}\left[1 - e_i \Big| Z_i = 1\right] + \frac{1}{N} \sum_{i \in [N]} L\mathbb{E}\left[e_i - (-1) \Big| Z_i = -1\right] \\
&\quad = \frac{2}{N} \frac{K}{K-1} L \sum_{i \in [N]} \sum_{j \notin \mathcal{C}(i)} \sum_s \frac{w_{is}}{\sum_s w_{is}} \frac{w_{js}}{\sum_k w_{ks}}.
\end{aligned}$$

We have now shown an upper bound on the argument of the right hand side of Equation (D.11). Next, we will show that $\tilde{g}_i(Z, e) := L \cdot e \; \forall i$ achieves this bound (and is therefore a maximizer over the class of L-Lipschitz functions) by directly computing the argument from the right hand side of Equation (D.11) under this choice of g. We use the definition of $\tilde{g}$ to bound the difference in $\tilde{g}$ by the difference in e, the fact that $e_i \in [-1, 1]$ to remove the absolute values, and finally apply Lemma D.2.2 with $\gamma_i = L$.

$$\begin{aligned}
&\left| \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[\tilde{g}_i(1,1) - \tilde{g}_i(1,e_i) | Z_i = 1\right] + \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[\tilde{g}_i(-1,e_i) - \tilde{g}_i(-1,-1) | Z_i = -1\right] \right| \qquad \text{(D.11)} \\
&\quad = \left| \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[L(1 - e_i) | Z_i = 1\right] + \frac{1}{N} \sum_{i \in [N]} \mathbb{E}\left[L(e_i - (-1)) | Z_i = -1\right] \right| \\
&\quad = \frac{1}{N} \sum_{i \in [N]} L\mathbb{E}\left[(1 - e_i) | Z_i = 1\right] + \frac{1}{N} \sum_{i \in [N]} L\mathbb{E}\left[(e_i - (-1)) | Z_i = -1\right] \\
&\quad = \frac{2}{N} \frac{K}{K-1} L \sum_{i \in [N]} \sum_{j \notin \mathcal{C}(i)} \sum_s \frac{w_{is}}{\sum_s w_{is}} \frac{w_{js}}{\sum_k w_{ks}}.
\end{aligned}$$

We conclude that $g_i(e) = L \cdot e \; \forall i$ achieves the maximum over L-Lipschitz functions g in Equation (D.11). Returning to that statement, we are now able to bound the bias using Equation (D.12) as

$$\arg\min_{\mathcal{C}} \max_{\{g_i \in \text{Lip}_L(e)\}} |\mathbb{E}[\hat{\tau}] - \tau^*| = \arg\min_{\mathcal{C}} \frac{2}{N} \frac{K}{K-1} L \sum_{i \in [N]} \sum_{j \notin \mathcal{C}(i)} \sum_s \frac{w_{is}}{\sum_s w_{is}} \frac{w_{js}}{\sum_k w_{ks}}.$$

Since K, N and L are constants with respect to the clustering $\mathcal{C}$, we recover the minimax optimality of the folded graph clustering:

$$\arg\min_{\mathcal{C}} \max_{\{g_i \in \text{Lip}_L(e)\}} |\mathbb{E}[\hat{\tau}] - \tau^*| = \arg\min_{\mathcal{C}} H(\mathcal{C})$$

D.2.4 Proof of Lemma 5.4.2 (bounding the bias under the $\Delta-$neighborhood potential outcomes model)

We begin by applying Lemma D.2.1 to decompose the bias in terms of its unit-level contributions:

$$\mathbb{E}[\hat{\tau}] - \tau^* = \frac{1}{N}\sum_{i\in[N]} \mathbb{E}\left[g_i(Z_i, e_i) - g(Z_i, Z_i)|Z_i = 1\right] - \frac{1}{N}\sum_{i\in[N]} \mathbb{E}\left[g_i(Z_i, e_i) - g(Z_i, Z_i)|Z_i = -1\right].$$

Next, we apply the property of the potential outcomes given in (5.4). We know from the assumption that whenever $|Z - e| < \Delta$, we have $|g_i(Z, e) - g_i(Z, Z)| = 0$, and otherwise we have $|g_i(Z, e) - g_i(Z, Z)| \leq B$. We can therefore bound the absolute value of the bias by m times the probability that each unit's exposure e_i deviates by more than Δ from its assignment Z_i.

$$\begin{aligned}
|\mathbb{E}[\hat{\tau}] - \tau^*| &\leq \frac{1}{N}\sum_{i\in[N]} \mathbb{E}\left[B\mathbf{1}\{e_i < 1 - \Delta\}|Z_i = 1\right] - \frac{1}{N}\sum_{i\in[N]} \mathbb{E}\left[B\mathbf{1}\{e_i > -1 + \Delta\}|Z_i = -1\right] \\
&= \frac{1}{N}\sum_{i\in[N]} B\mathbb{P}\left(e_i < 1 - \Delta|Z_i = 1\right) - \frac{1}{N}\sum_{i\in[N]} B\mathbb{P}\left(\mathbf{1}\{e_i > -1 + \Delta\}|Z_i = -1\right) \\
&= \frac{1}{N}\sum_{i\in[N]} B\mathbb{P}\left(1 - e_i > \Delta|Z_i = 1\right) - \frac{1}{N}\sum_{i\in[N]} B\mathbb{P}\left(1 + e_i > \Delta|Z_i = -1\right)
\end{aligned}$$

We apply Markov's inequality to both terms:

$$|\mathbb{E}[\hat{\tau}] - \tau^*| \leq \sum_{i\in[N]} \frac{B}{\Delta}\left(\mathbb{E}[1 - e_i|Z_i = 1] + \mathbb{E}[1 + e_i|Z_i = -1]\right)$$

Next, we apply Lemma D.2.2 with $\gamma_i = 1$ to write this expression in terms of our folded graph clustering objective, which completes the proof.

$$|\mathbb{E}[\hat{\tau}] - \tau^*| \leq \frac{2B}{N\Delta}\frac{K}{K-1}\sum_{i\in[N]}\sum_{j\notin\mathcal{C}(i)}\sum_{s}\frac{w_{is}}{\sum_s w_{is}}\frac{w_{js}}{\sum_k w_{ks}}.$$

□

D.2.5 Proof of Lemma 5.3.3 (The objective $\mathcal{H}(\mathcal{C})$ maximizes the covariance between exposure and treatment assignment)

In this proof, we will use the linearity of $\mathbf{e}$ in $\mathbf{Z}$ to write the covariance objective entirely in terms of linear combinations of Z_i. We will then use the fact that $\mathcal{D}(\mathcal{C})$ is a cluster-randomized design to compute the covariance exactly, and show that minimizing the covariance objective is identical to minimizing the folded graph clustering objective.

We begin by rewriting our optimization objective in terms of only the treatment assignments Z. Recall that, under the linear dose and exposure mappings, Lemma D.2.3, $\mathbf{e} = C\mathbf{Z}$ for a known matrix C that depends only on the interference graph.

$$\begin{aligned}\arg\max_{\mathcal{C}}\ Tr\left(Cov_{Z\sim\mathcal{D}(\mathcal{C})}(\mathbf{Z},\mathbf{e})\right) &= \arg\max_{\mathcal{C}}\ Tr\left(\mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}\left[(\mathbf{Z}-\mathbb{E}[\mathbf{Z}])(\mathbf{e}-\mathbb{E}[\mathbf{e}])^T\right]\right)\\ &= \arg\max_{\mathcal{C}}\ Tr\left(\mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}\left[(\mathbf{Z}-\mathbb{E}[\mathbf{Z}])(C\mathbf{Z}-\mathbb{E}[C\mathbf{Z}])^T\right]\right)\\ &= \arg\max_{\mathcal{C}}\ \mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}\left[Tr\left(C(\mathbf{Z}-\mathbb{E}[\mathbf{Z}])(\mathbf{Z}-\mathbb{E}[\mathbf{Z}])^T\right)\right]\end{aligned}$$

Observe that we are taking the trace of a product of two $N \times N$ matrices, C and $(\mathbf{Z}-\mathbb{E}[\mathbf{Z}])(\mathbf{Z}-\mathbb{E}[\mathbf{Z}])^T$. The trace of a product of square matrices is the sum of the entries in their elementwise (Hadamard) product, which lets us write the trace as a sum and apply linearity of expectation:

$$\begin{aligned}\arg\max_{\mathcal{C}}\ Tr\left(Cov_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}(\mathbf{e},\mathbf{Z})\right) &= \arg\max_{\mathcal{C}}\ \mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}\left[\sum_{i,j\in[N]}(Z_i-\mathbb{E}[Z_i])(Z_j-\mathbb{E}[Z_j])C_{ij}\right]\\ &= \arg\max_{\mathcal{C}}\ \sum_{i,j\in[N]}C_{ij}\mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}\left[(Z_i-\mathbb{E}[Z_i])(Z_j-\mathbb{E}[Z_j])\right]\end{aligned} \tag{D.12}$$

We have by Definition 5.3.1 that $\mathcal{D}(\mathcal{C})$ is a uniform assignment of K balanced clusters into K_T treated units and $K_C = K - K_T$ control units. We see that the value of the expectation $\mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}\left[(Z_i-\mathbb{E}[Z_i])(Z_j-\mathbb{E}[Z_j])\right]$ depends on whether units i and j belong to the same or different clusters. We have

$$\begin{aligned}\arg\max_{\mathcal{C}}\ Tr\left(Cov_{Z\sim\mathcal{D}(\mathcal{C})}(\mathbf{e},\mathbf{Z})\right) = \arg\max_{\mathcal{C}}\ \sum_i\Bigg(&\sum_{j\in\mathcal{C}(i)}C_{ij}\mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}\left[(Z_i-\mathbb{E}[Z_i])(Z_j-\mathbb{E}[Z_j])|j\in\mathcal{C}(i)\right]\\ &+\sum_{j\notin\mathcal{C}(i)}C_{ij}\mathbb{E}_{Z\sim\mathcal{D}}\left[(Z_i-\mathbb{E}[Z_i])(Z_j-\mathbb{E}[Z_j])|j\notin\mathcal{C}(i)\right]\Bigg)\end{aligned}$$

Our next step will be to compute both conditional expectations. We will see that the conditional expectations are independent of the indices i and j (since the argument of the expectation depends only on whether i and j belong to the same cluster), and that the conditional expectation is greater when i and j belong to different clusters. This will let us draw an equivalence between our covariance objective and the objective of minimizing the folded graph cut, which will turn out to be exactly the objective of minimizing cuts in C_{ij}. We expand the conditional covariances and

use linearity of expectation, along with the fact that $\mathbb{E}[Z_j] = (K_T - K_C)/K$ to write

$$\begin{aligned}
&\mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}\left[(Z_i - \mathbb{E}[Z_i])(Z_j - \mathbb{E}[Z_j])|j \in \mathcal{C}(i)\right] \\
&\quad= \mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}\left[Z_i Z_j|j \in \mathcal{C}(i)\right] - \mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}\left[Z_i|j \in \mathcal{C}(i)\right]\mathbb{E}[Z_j] - \\
&\qquad\qquad \mathbb{E}_{Z\sim\mathcal{D}(\mathcal{C})}\left[Z_j|j \in \mathcal{C}(i)\right]\mathbb{E}[Z_i] + \mathbb{E}[Z_j]\mathbb{E}[Z_i] \\
&\quad= \mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}\left[Z_i Z_j|j \in \mathcal{C}(i)\right] - \mathbb{E}\left[Z_i\right]\mathbb{E}[Z_j] - \mathbb{E}\left[Z_j\right]\mathbb{E}[Z_i] + \mathbb{E}[Z_j]\mathbb{E}[Z_i] \\
&\quad= \mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}\left[Z_i Z_j|j \in \mathcal{C}(i)\right] - \mathbb{E}[Z_j]\mathbb{E}[Z_i] \\
&\quad= \mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}\left[Z_i Z_j|j \in \mathcal{C}(i)\right] - \left(\frac{K_T - K_C}{K}\right)^2 \\
&\quad= 1 - \left(\frac{K_T - K_C}{K}\right)^2
\end{aligned}$$

and similarly

$$\begin{aligned}
&\mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}\left[(Z_i - \mathbb{E}[Z_i])(Z_j - \mathbb{E}[Z_j])|j \notin \mathcal{C}(i)\right] \\
&\quad= \mathbb{E}_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}\left[Z_i Z_j|j \notin \mathcal{C}(i)\right] - \mathbb{E}[Z_j]\mathbb{E}[Z_i] \\
&\quad= 1 \cdot \frac{K_T}{K} \cdot \frac{K_T - 1}{K - 1} + (-1) \cdot \frac{K_T}{K}\frac{K_C - 1}{K - 1} + \\
&\qquad\qquad (-1) \cdot \frac{K_C}{K}\frac{K_T - 1}{K - 1} + 1 \cdot \frac{K_C}{K}\frac{K_C - 1}{K - 1} - \mathbb{E}[Z_j]\mathbb{E}[Z_i] \\
&\quad= \frac{K_T(K_T - 1) - K_T(K_C - 1) - K_C(K_T - 1) + K_C(K_C - 1)}{K(K - 1)} - \mathbb{E}[Z_j]\mathbb{E}[Z_i] \\
&\quad= \frac{K_T(K_T - 1) - K_T(K_C - 1) - K_C(K_T - 1) + K_C(K_C - 1)}{K(K - 1)} - \left(\frac{K_T - K_C}{K}\right)^2 \\
&\quad= \frac{(K_T - K_C)^2}{K(K - 1)} - \left(\frac{K_T - K_C}{K}\right)^2
\end{aligned}$$

Substituting these conditional expectations, we have

$$\begin{aligned}
\arg\max_{\mathcal{C}}\ Tr\left(Cov_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}(\mathbf{e}, \mathbf{Z})\right) = \arg\max_{\mathcal{C}}\ \sum_i \Bigg(&\sum_{j\in[N]} C_{ij}\left(1 - \left(\frac{K_T - K_C}{K}\right)^2\right) + \\
&\sum_{j\notin\mathcal{C}(i)} C_{ij}\left(\frac{(K_T - K_C)^2}{K(K - 1)} - 1\right)\Bigg)
\end{aligned}$$

Next, we recognize that the sum over all i and j is constant with respect to the cluster randomized

design $\mathcal{C}$ and therefore can be removed from the argmax.

$$\begin{aligned}\arg\max_{\mathcal{C}} \; Tr\left(Cov_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}(\mathbf{e},\mathbf{Z})\right) &= \arg\max_{\mathcal{C}} \; \sum_i \left(\sum_{j\notin\mathcal{C}(i)} C_{ij}\left(\frac{(K_T-K_C)^2}{K(K-1)}-1\right)\right)\\ &= \arg\min_{\mathcal{C}} \; \sum_i \left(\sum_{j\notin\mathcal{C}(i)} C_{ij}\left(1-\frac{(K_T-K_C)^2}{K(K-1)}\right)\right)\end{aligned}$$

Observe that, as long as there is at least one treated and control unit (i.e., $0 < K_T < K$) then the quantity $1-\frac{(K_T-K_C)^2}{K(K-1)}$ is positive and can be taken out of the argmax. In this case, we have

$$\arg\max_{\mathcal{C}} \; Tr\left(Cov_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}(\mathbf{e},\mathbf{Z})\right) = \arg\min_{\mathcal{C}} \; \sum_i \left(\sum_{j\notin\mathcal{C}(i)} C_{ij}\right)$$

Finally, we substitute the value of C_{ij} from Lemma D.2.3:

$$\arg\max_{\mathcal{C}} \; Tr\left(Cov_{\mathbf{Z}\sim\mathcal{D}(\mathcal{C})}(\mathbf{e},\mathbf{Z})\right) = \arg\min_{\mathcal{C}} \; \sum_i \left(\sum_{j\notin\mathcal{C}(i)} \frac{\sum_s w_{is}\frac{w_{js}}{\sum_k w_{ks}}}{\sum_s w_{is}}\right)$$

which we recognize as precisely the minimum-cut objective on the folded graph. □

D.3 Standard deviation and RMSE for experiments

We set up the parameters of our first two experiments so that the error of $\widehat{\tau}_{DIM}$ was almost entirely due to bias, instead of variance. Here we provide the normalized standard deviation and RMSE for the experiments in Sections 5.5.1 and 5.5.2, for completeness.

D.3.1 Robustness to different graph structures

Tables D.3 and D.4 provide the normalized standard deviation and RMSE for the experiments in Sections 5.5.1.

D.3.2 Robustness to nonlinearity

Tables D.5 and D.6 provide the normalized standard deviation and RMSE for the experiments in Sections 5.5.2.

Table D.3: Relative RMSE of $\widehat{\tau}_{DIM}$ as the bipartite stochastic block model changes (see 5.5.1)

	$p = 0.0$	$p = 0.005$	$p = 0.05$	$p = 0.5$
$\mathcal{H}(\mathcal{C})$	0.23($\pm$0.03)	3.84($\pm$0.05)	11.54($\pm$0.06)	12.99($\pm$0.09)
Tr(Var($\mathbf{d}$))	0.28($\pm$0.03)	3.86($\pm$0.05)	11.49($\pm$0.06)	12.92($\pm$0.07)
Direct clustering	0.26($\pm$0.03)	9.26($\pm$0.14)	12.69($\pm$0.07)	12.95($\pm$0.07)
EXPOSURE-DESIGN	0.53($\pm$0.05)	4.06($\pm$0.07)	11.9($\pm$0.07)	13.0($\pm$0.08)
Unit-level randomization	12.43($\pm$0.08)	12.55($\pm$0.09)	12.77($\pm$0.07)	12.96($\pm$0.08)
True clusters	0.31($\pm$0.04)	3.86($\pm$0.04)	11.58($\pm$0.06)	12.96($\pm$0.06)

Table D.4: Relative standard deviation of $\widehat{\tau}_{DIM}$ as the bipartite stochastic block model changes (see 5.5.1)

	$p = 0.0$	$p = 0.005$	$p = 0.05$	$p = 0.5$
$\mathcal{H}(\mathcal{C})$	0.23($\pm$0.03)	0.26($\pm$0.04)	0.28($\pm$0.04)	0.42($\pm$0.05)
Tr(Var($\mathbf{d}$))	0.28($\pm$0.04)	0.28($\pm$0.03)	0.31($\pm$0.05)	0.37($\pm$0.05)
Direct clustering	0.26($\pm$0.03)	0.69($\pm$0.09)	0.37($\pm$0.05)	0.35($\pm$0.05)
EXPOSURE-DESIGN	0.37($\pm$0.04)	0.33($\pm$0.04)	0.33($\pm$0.04)	0.41($\pm$0.05)
Unit-level Randomization	0.39($\pm$0.05)	0.45($\pm$0.07)	0.38($\pm$0.05)	0.41($\pm$0.06)
True Clusters	0.31($\pm$0.03)	0.23($\pm$0.03)	0.28($\pm$0.03)	0.29($\pm$0.04)

Table D.5: Relative RMSE of $\widehat{\tau}_{DIM}$ as the neighborhood of pure exposure, Δ, widens (see 5.5.2)

	$\Delta = 0.1$	$\Delta = 0.3$	$\Delta = 0.5$
$\mathcal{H}(\mathcal{C})$	1.0($\pm$0.004)	0.458($\pm$0.005)	0.001($\pm$0.0)
Tr(Var($\mathbf{d}$))	1.002($\pm$0.004)	0.461($\pm$0.004)	0.001($\pm$0.0)
Direct clustering	0.997($\pm$0.005)	0.95($\pm$0.008)	0.608($\pm$0.02)
EXPOSURE-DESIGN	1.001($\pm$0.004)	0.509($\pm$0.005)	0.011($\pm$0.002)
Unit-level randomization	0.998($\pm$0.004)	1.0($\pm$0.004)	0.998($\pm$0.003)
True clusters	1.001($\pm$0.004)	0.459($\pm$0.004)	0.001($\pm$0.0)

Table D.6: Relative standard deviation of $\widehat{\tau}_{DIM}$ as the neighborhood of pure exposure, Δ, widens (see 5.5.2)

	$\Delta = 0.1$	$\Delta = 0.3$	$\Delta = 0.5$
$\mathcal{H}(\mathcal{C})$	0.02($\pm$0.003)	0.024($\pm$0.003)	0.001($\pm$0.0)
$\text{Tr}(\text{Var}(\mathbf{d}))$	0.02($\pm$0.002)	0.021($\pm$0.003)	0.001($\pm$0.0)
Direct clustering	0.026($\pm$0.003)	0.038($\pm$0.007)	0.097($\pm$0.01)
EXPOSURE-DESIGN	0.02($\pm$0.003)	0.025($\pm$0.003)	0.005($\pm$0.001)
Unit-level randomization	0.019($\pm$0.002)	0.019($\pm$0.002)	0.017($\pm$0.002)
True clusters	0.021($\pm$0.003)	0.024($\pm$0.003)	0.001($\pm$0.0)

Bibliography

The European committee on antimicrobial susceptibility testing breakpoint tables for interpretation of mics and zone diameters. version 11.0, 2021.

Akshay Agrawal, Robin Verschueren, Steven Diamond, and Stephen Boyd. A rewriting system for convex optimization problems. *Journal of Control and Decision*, 5(1):42–60, 2018.

Truven Health Analytics. United States marketscan commercial claims and encounters database 2010, 2010.

Konstantin Andreev and Harald Racke. Balanced graph partitioning. *Theory of Computing Systems*, 39(6):929–939, 2006.

Sinan Aral and Dylan Walker. Creating social contagion through viral product design: A randomized trial of peer influence in networks. *Management science*, 57(9):1623–1639, 2011.

Peter M Aronow. A general method for detecting interference between units in randomized experiments. *Sociological Methods & Research*, 41(1):3–16, 2012.

Peter M Aronow and Cyrus Samii. Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics*, 11(4):1912–1947, 2017.

Peter M Aronow, Dean Eckles, Cyrus Samii, and Stephanie Zonszein. Spillover effects in experimental data. *Advances in Experimental Political Science*, 289:319, 2021.

Susan Athey, Dean Eckles, and Guido W Imbens. Exact p-values for network interference. *Journal of the American Statistical Association*, 113(521):230–240, 2018.

Kevin Aydin, MohammadHossein Bateni, and Vahab Mirrokni. Distributed balanced partitioning via linear embedding. *Algorithms*, 12(8):162, 2019.

Patrick Bajari, Brian Burdick, Guido W Imbens, Lorenzo Masoero, James McQueen, Thomas Richardson, and Ido M Rosen. Multiple randomization designs. *arXiv preprint arXiv:2112.13495*, 2021.

Guillaume W Basse and Edoardo M Airoldi. Limitations of design-based causal inference and A/B testing under arbitrary and network interference. *Sociological Methodology*, 48(1):136–151, 2018.

Mikhail Belkin and Kaushik Sinha. Polynomial learning of distribution families. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 103–112. IEEE, 2010.

Yoav Benjamini and Yosef Hochberg. On the adaptive control of the false discovery rate in multiple testing with independent statistics. *Journal of educational and Behavioral Statistics*, 25(1):60–83, 2000.

Casey Beppler, Elif Tekin, Zhiyuan Mao, Cynthia White, Cassandra McDiarmid, Emily Vargas, Jeffrey H Miller, Van M Savage, and Pamela J Yeh. Uncovering emergent interactions in three-way combinations of stressors. *Journal of the Royal Society Interface*, 13(125):20160800, 2016.

Casey Beppler, Elif Tekin, Cynthia White, Zhiyuan Mao, Jeffrey H Miller, Robert Damoiseaux, Van M Savage, and Pamela J Yeh. When more is less: emergent suppressive interactions in three-drug combinations. *BMC microbiology*, 17(1):1–9, 2017.

Morris C Berenbaum. Synergy, additivism and antagonism in immunosuppression. A critical review. *Clinical and experimental immunology*, 28(1):1, 1977.

Morris C Berenbaum. What is synergy? *Pharmacological reviews*, 41(2):93–141, 1989.

Thomas Blake and Dominic Coey. Why marketplace experimentation is harder than it seems: The role of test-control interference. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 567–582, 2014.

Chester I Bliss. The toxicity of poisons applied jointly. *Annals of applied biology*, 26(3):585–615, 1939.

Jane H Booth, Shalom I Benrimoj, and Graeme R Nimmo. In vitro interactions of neomycin sulfate, bacitracin, and polymyxin B sulfate. *International journal of dermatology*, 33(7):517–520, 1994.

Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

Jennifer Brennan, Ramya Korlakai Vinayak, and Kevin Jamieson. Estimating the number and effect sizes of non-null hypotheses. In *International Conference on Machine Learning*, pages 1123–1133. PMLR, 2020.

Jennifer Brennan, Marlena Bannick, Nicholas Kassebaum, Lauren Wilner, Azalea Thomson, Aleksandr Aravkin, and Peng Zheng. Analysis and methods to mitigate effects of under-reporting in count data. *arXiv preprint arXiv:2109.12247*, 2021.

Jennifer Brennan, Lalit Jain, Sofia Garman, Ann E Donnelly, Erik Scott Wright, and Kevin Jamieson. Sample-efficient identification of high-dimensional antibiotic synergy with a normalized diagonal sampling design. *PLOS Computational Biology*, 18(7):e1010311, 2022.

SRM Bushby and GH Hitchings. Trimethoprim, a sulphonamide potentiator. *British journal of pharmacology and chemotherapy*, 33(1):72–90, 1968.

T Tony Cai, Jiashun Jin, Mark G Low, et al. Estimation and confidence sets for sparse normal mixtures. *The Annals of Statistics*, 35(6):2421–2449, 2007.

Ozan Candogan, Chen Chen, and Rad Niazadeh. Near-optimal experimental design for networks: Independent block randomization. *Chicago Booth Research Paper*, (21-17), 2021.

Alexandra Carpentier and Nicolas Verzelen. Adaptive estimation of the sparsity in the Gaussian vector model. *The Annals of Statistics*, 47(1):93–126, 2019.

Nina Cedergreen. Quantifying synergy: a systematic review of mixture toxicity studies within environmental toxicology. *PloS one*, 9(5):e96580, 2014.

N Chamandy. Experimentation in a ridesharing marketplace. *URL: https://eng.lyft.com/experimentation-in-a-ridesharing-marketplace-b39db027a66e*, 2016.

Karthekeyan Chandrasekaran and Richard Karp. Finding a most biased coin with fewest flips. In *Conference on Learning Theory*, pages 394–407, 2014.

Sriram Chandrasekaran, Melike Cokol-Cakmak, Nil Sahin, Kaan Yilancioglu, Hilal Kazan, James J Collins, and Murat Cokol. Chemogenomics and orthology-based design of antibiotic combination therapies. *Molecular systems biology*, 12(5):872, 2016.

Xiongzhi Chen. Uniformly consistently estimating the proportion of false null hypotheses via Lebesgue–Stieltjes integral equations. *Journal of Multivariate Analysis*, 2019.

Ting-Chao Chou. Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. *Pharmacological reviews*, 58(3):621–681, 2006.

Ting-Chao Chou and Paul Talalay. Analysis of combined drug effects: a new look at a very old problem. *Trends in Pharmacological Sciences*, 4:450–454, 1983.

Murat Cokol, Nurdan Kuru, Ece Bicak, Jonah Larkins-Ford, and Bree B Aldridge. Efficient measurement and factorization of high-order drug interactions in Mycobacterium tuberculosis. *Science advances*, 3(10):e1701881, 2017.

Murat Cokol, Chen Li, and Sriram Chandrasekaran. Chemogenomic model identifies synergistic drug combinations robust to the pathogen microenvironment. *PLoS computational biology*, 14 (12):e1006677, 2018.

Melike Cokol-Cakmak, Selim Cetiner, Nurdan Erdem, Feray Bakan, and Murat Cokol. Guided screen for synergistic three-drug combinations. *PloS one*, 15(7):e0235929, 2020.

Constantinos Daskalakis and Gautam Kamath. Faster and sample near-optimal algorithms for proper learning mixtures of Gaussians. In *Conference on Learning Theory*, pages 1183–1213, 2014.

Carl De Boor. *A practical guide to splines*, volume 27. Springer-Verlag New York, 1978.

Steven Diamond and Stephen Boyd. CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 17(83):1–5, 2016.

Peng Ding, Xinran Li, and Luke W Miratrix. Bridging finite and super population causal inference. *Journal of Causal Inference*, 5(2), 2017.

David Donoho and Jiashun Jin. Higher criticism for detecting sparse heterogeneous mixtures. *Ann. Statist.*, 32(3):962–994, 06 2004. doi: 10.1214/009053604000000265.

Olive Jean Dunn. Multiple comparisons among means. *Journal of the American statistical association*, 56(293):52–64, 1961.

Michaela Dvorzak and Helga Wagner. Sparse Bayesian modelling of underreported count data. *Statistical Modelling*, 16(1):24–46, 2016.

Harry Eagle and AD Musselman. The rate of bactericidal action of penicillin in vitro as a function of its concentration, and its paradoxically reduced activity at high concentrations against certain organisms. *The Journal of experimental medicine*, 88(1):99–131, 1948.

Dean Eckles, René F Kizilcec, and Eytan Bakshy. Estimating peer effects in networks with peer encouragement designs. *Proceedings of the National Academy of Sciences*, 113(27):7316–7322, 2016.

Dean Eckles, Brian Karrer, and Johan Ugander. Design and analysis of experiments in networks: Reducing bias from interference. *Journal of Causal Inference*, 5(1), 2017.

Bradley Efron et al. Size, power and false discovery rates. *The Annals of Statistics*, 35(4):1351–1377, 2007.

Peter Fader and Bruce Hardie. A note on modelling underreported Poisson counts. *Journal of Applied Statistics*, 27(8):953–964, 2000.

Julie Foucquier and Mickael Guedj. Analysis of drug combinations: Current methodological landscape. *Pharmacology research & perspectives*, 3(3):e00149, 2015.

Christopher Genovese, Larry Wasserman, et al. A stochastic process approach to false discovery control. *The Annals of Statistics*, 32(3):1035–1061, 2004.

Huan Gui, Ya Xu, Anmol Bhasin, and Jiawei Han. Network A/B testing: From sampling to estimation. In *Proceedings of the 24th International Conference on World Wide Web*, pages 399–409, 2015.

MJ Hall, RF Middleton, and D Westmacott. The fractional inhibitory concentration (FIC) index as a measure of synergy. *Journal of Antimicrobial Chemotherapy*, 11(5):427–433, 1983.

Linhui Hao, Akira Sakurai, Tokiko Watanabe, Ericka Sorensen, Chairul A Nidom, Michael A Newton, Paul Ahlquist, and Yoshihiro Kawaoka. Drosophila RNAi screen identifies host genes important for influenza virus replication. *Nature*, 454(7206):890, 2008.

Moritz Hardt and Eric Price. Tight bounds for learning a mixture of two Gaussians. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 753–760. ACM, 2015.

Christopher Harshaw, Fredrik Sävje, David Eisenstat, Vahab Mirrokni, and Jean Pouget-Abadie. Design and analysis of bipartite experiments under a linear exposure-response model. *Proceedings of the 23rd ACM Conference on Economics and Computation (EC '22)*, 2022.

Jan FA Hendrickx, Edmond I Eger, James M Sonner, and Steven L Shafer. Is synergy the rule? a review of anesthetic interactions producing hypnosis and immobility. *Anesthesia & Analgesia*, 107(2):494–506, 2008.

Arthur E Hoerl and Robert W Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970.

Michael G Hudgens and M Elizabeth Halloran. Toward causal inference with interference. *Journal of the American Statistical Association*, 103(482):832–842, 2008.

Guido W Imbens and Donald B Rubin. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press, 2015.

Kevin G Jamieson and Lalit Jain. A bandit approach to sequential experimental design with false discovery control. In *Advances in Neural Information Processing Systems 31*, pages 3660–3670. 2018.

Kevin G Jamieson, Daniel Haas, and Benjamin Recht. The power of adaptivity in identifying statistical alternatives. In *Advances in Neural Information Processing Systems*, pages 775–783, 2016.

Wenhua Jiang and Cun-Hui Zhang. Generalized likelihood ratio test for normal mixtures. *Statistica Sinica*, 26:955–978, 07 2016. doi: 10.5705/ss.202015.0086.

Jiashun Jin. Proportion of non-zero normal means: Universal oracle equivalences and uniformly consistent estimators. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(3):461–493, 2008.

Ramesh Johari, Hannah Li, Inessa Liskovich, and Gabriel Y Weintraub. Experimental design in two-sided platforms: An analysis of bias. *Management Science*, 2022.

Adam Tauman Kalai, Ankur Moitra, and Gregory Valiant. Efficiently learning mixtures of two Gaussians. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 553–562. ACM, 2010.

Itay Katzir, Murat Cokol, Bree B Aldridge, and Uri Alon. Prediction of ultra-high-order antibiotic combinations based on pairwise interactions. *PLoS computational biology*, 15(1):e1006774, 2019.

Göran Kauermann and Raymond J Carroll. A note on the efficiency of sandwich covariance matrix estimation. *Journal of the American Statistical Association*, 96(456):1387–1396, 2001.

Bor Kavčič, Gašper Tkačik, and Tobias Bollenbach. Minimal biophysical model of combined antibiotic action. *PLoS Computational Biology*, 17(1):e1008529, 2021.

Marisa Kirisame, Steven Lyubomirsky, Altan Haan, Jennifer Brennan, Mike He, Jared Roesch, Tianqi Chen, and Zachary Tatlock. Dynamic tensor rematerialization. In *International Conference on Learning Representations*, 2021.

Kasper Nørskov Kragh, Desiree Gijón, Ainhize Maruri, Alberto Antonelli, Marco Coppi, Mette Kolpen, Stephanie Crone, Chaitanya Tellapragada, Badrul Hasan, Stine Radmer, et al. Effective antimicrobial combination in vivo treatment predicted with microcalorimetry screening. *Journal of Antimicrobial Chemotherapy*, 76(4):1001–1009, 2021.

Stephen E Kurtz, Christopher A Eide, Andy Kaempf, Motomi Mori, Cristina E Tognon, Uma Borate, Brian J Druker, and Jeffrey W Tyner. Dual inhibition of jak1/2 kinases and bcl2: a promising therapeutic strategy for acute myeloid leukemia. *Leukemia*, 32(9):2025–2028, 2018.

Jasper CH Lee and Paul Valiant. Uncertainty about uncertainty: Optimal adaptive algorithms for estimating mixtures of unknown coins. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 414–433. SIAM, 2021.

Ang Li and Rina Foygel Barber. Multiple testing with the structure-adaptive Benjamini–Hochberg algorithm. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 81(1):45–74, 2019.

Hannah Li, Geng Zhao, Ramesh Johari, and Gabriel Y Weintraub. Interference, bias, and variance in two-sided marketplace experimentation: Guidance for platforms. In *Proceedings of the ACM Web Conference 2022*, pages 182–192, 2022.

S Loewe. The problem of synergism and antagonism of combined drugs. *Arzneimittel Forschung*, 3 (6):289–90, June 1953.

Martin Lukačišin and Tobias Bollenbach. Emergent gene expression responses to drug combinations predict higher-order drug interactions. *Cell Systems*, 9(5):423–433, 2019.

Charles F Manski. Identification of treatment response with social interactions. *The Econometrics Journal*, 16(1):S1–S23, 2013.

Pascal Massart. The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. *The annals of Probability*, pages 1269–1283, 1990.

Nicolai Meinshausen and Peter Bühlmann. Lower bounds for the number of false null hypotheses for multiple testing of associations under general dependence structures. *Biometrika*, 92(4): 893–907, 2005.

Nicolai Meinshausen and John Rice. Estimating the proportion of false null hypotheses among a large number of independently tested hypotheses. *The Annals of Statistics*, 34(1):373–393, 2006.

Ankur Moitra and Gregory Valiant. Settling the polynomial learnability of mixtures of Gaussians. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 93–102. IEEE, 2010.

Evan Munro, Stefan Wager, and Kuang Xu. Treatment effects in market equilibrium. *arXiv preprint arXiv:2109.11647*, 2021.

Josue Nassar, Jennifer Brennan, Ben Evans, and Kendall Lowrey. BAM: Bayes with adaptive memory. In *International Conference on Learning Representations*, 2022.

Centers for Disease Control National Center for Health Statistics and Prevention. United States national ambulatory medical care survey, 2018.

Centers for Disease Control National Center for Health Statistics and United States Census Bureau (USCB) Prevention. United States national hospital discharge survey, 2010.

Jerzy Neyman. On the application of probability theory to agricultural experiments. Essay on principles. Section 9 (translated). Reprinted ed. *Statistical Science*, 5:465–472, 1923.

Joel Nishimura and Johan Ugander. Restreaming graph partitioning: simple versatile algorithms for advanced balancing. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1106–1114, 2013.

Jonas Noeske, Jian Huang, Nelson B Olivier, Robert A Giacobbe, Mark Zambrowski, and Jamie HD Cate. Synergy of streptogramin antibiotics occurs independently of their effects on translation. *Antimicrobial agents and chemotherapy*, 58(9):5269–5279, 2014.

Jennifer O'Neil, Yair Benita, Igor Feldman, Melissa Chenard, Brian Roberts, Yaping Liu, Jing Li, Astrid Kral, Serguei Lejnine, Andrey Loboda, et al. An unbiased oncology compound screen to identify novel combination strategies. *Molecular cancer therapeutics*, 15(6):1155–1162, 2016.

Georgios Papadopoulos and Joao Santos Silva. Identification issues in models for underreported counts. 2008.

Georgios Papadopoulos and J.M.C. Santos Silva. Identification issues in some double-index models for non-negative data. *Economics Letters*, 117(1):365–367, oct 2012. doi: 10.1016/j.econlet.2012.06.001.

Giovanni Parmigiani and Lurdes Inoue. *Decision theory: Principles and approaches*, volume 812. John Wiley & Sons, 2009.

Rohit Kumar Patra and Bodhisattva Sen. Estimation of a two-component mixture model with applications to multiple testing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 78(4):869–893, 2016.

Karl Pearson. Contributions to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society of London. A*, 185:71–110, 1894.

Jean Pouget-Abadie, Kevin Aydin, Warren Schudy, Kay Brodersen, and Vahab Mirrokni. Variance reduction in bipartite experiments through correlation clustering. *Advances in Neural Information Processing Systems*, 32:13309–13319, 2019.

Anggia Prasetyoputri, Angie M Jarrad, Matthew A Cooper, and Mark AT Blaskovich. The Eagle effect and antibiotic-induced persistence: Two sides of the same coin? *Trends in microbiology*, 27(4):339–354, 2019.

Luc Pronzato and Andrej Pázman. Design of experiments in nonlinear models. *Lecture notes in statistics*, 212:1, 2013.

Friedrich Pukelsheim. *Optimal design of experiments*. SIAM, 2006.

David Rolnick, Kevin Aydin, Jean Pouget-Abadie, Shahab Kamali, Vahab Mirrokni, and Amir Najmi. Randomized experimental design via geographic clustering. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2745–2753, 2019.

Donald Rubin. Discussion of "Randomization analysis of experimental data in the Fisher randomization test" by D. Basu. *Journal of the American statistical association*, 75:591–593, 1980.

Donald B Rubin. Comment: Neyman (1923) and causal inference in experiments and observational studies. *Statistical Science*, 5(4):472–480, 1990.

D Russ and R Kishony. Additivity of inhibitory effects in multidrug combinations. *Nature microbiology*, 3(12):1339–1345, 2018.

Martin Saveski, Jean Pouget-Abadie, Guillaume Saint-Jacques, Weitao Duan, Souvik Ghosh, Ya Xu, and Edoardo M Airoldi. Detecting network effects: Randomizing over randomized experiments. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1027–1035, 2017.

Fredrik Sävje, Peter M Aronow, and Michael G Hudgens. Average treatment effects in the presence of unknown interference. *The Annals of Statistics*, 49(2):673–701, 2021.

David C Schmittlein, Albert C Bemmaor, and Donald G Morrison. Why does the NBD model work? Robustness in representing product purchases, brand purchases and imperfectly recorded purchases. *Marketing Science*, 4(3):255–266, 1985.

Tore Schweder and Eil Spjøtvoll. Plots of p-values to evaluate many tests simultaneously. *Biometrika*, 69(3):493–502, 1982.

Charity D Scripture and William D Figg. Drug interactions in cancer therapy. *Nature Reviews Cancer*, 6(7):546–558, 2006.

James D Stamey, Dean M Young, and Doyle Boese. A Bayesian hierarchical model for Poisson rate and reporting-probability inference using double sampling. *Australian & New Zealand Journal of Statistics*, 48(2):201–212, 2006.

Matthew Stephens. False discovery rates: a new deal. *Biostatistics*, 18(2):275–294, 2017.

Oliver Stoner, Theo Economou, and Gabriela Drummond Marques da Silva. A hierarchical framework for correcting under-reporting in count data. *Journal of the American Statistical Association*, 114(528):1481–1492, 2019.

John D Storey. A direct approach to false discovery rates. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(3):479–498, 2002.

Elif Tekin, Casey Beppler, Cynthia White, Zhiyuan Mao, Van M Savage, and Pamela J Yeh. Enhanced identification of synergistic and antagonistic emergent interactions among three or more drugs. *Journal of The Royal Society Interface*, 13(119):20160332, 2016.

Elif Tekin, Cynthia White, Tina Manzhu Kang, Nina Singh, Mauricio Cruz-Loya, Robert Damoiseaux, Van M Savage, and Pamela J Yeh. Prevalence and patterns of higher-order drug interactions in Escherichia coli. *NPJ systems biology and applications*, 4(1):1–10, 2018.

M'hamed Temkit. *Experimental Designs for Generalized Linear Models and Functional Magnetic Resonance Imaging*. Arizona State University, 2014. URL `https://hdl.handle.net/2286/R.I.27465`.

John Thickstun, Jennifer Brennan, and Harsh Verma. Rethinking evaluation methodology for audio-to-score alignment. *arXiv preprint arXiv:2009.14374*, 2020.

Kevin Tian, Weihao Kong, and Gregory Valiant. Learning populations of parameters. In *Advances in Neural Information Processing Systems*, pages 5778–5787, 2017.

Panos Toulis and Edward Kao. Estimation of causal peer influence effects. In *International conference on machine learning*, pages 1489–1497. PMLR, 2013.

Alexandre B. Tsybakov. *Introduction to Nonparametric Estimation*. Springer-Verlag New York, 1 edition, 2009. ISBN 978-0-387-79051-0.

Mike Tyers and Gerard D Wright. Drug combinations: a strategy to extend the life of antibiotics in the 21st century. *Nature Reviews Microbiology*, 17(3):141–155, 2019.

Johan Ugander, Brian Karrer, Lars Backstrom, and Jon Kleinberg. Graph cluster randomization: Network exposure to multiple universes. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 329–337, 2013.

Ramya Korlakai Vinayak, Weihao Kong, Gregory Valiant, and Sham Kakade. Maximum likelihood estimation for learning populations of parameters. In *International Conference on Machine Learning*, pages 6448–6457, 2019.

Theo Vos, Stephen S Lim, Cristiana Abbafati, Kaja M Abbas, Mohammad Abbasi, Mitra Abbasifard, Mohsen Abbasi-Kangevari, Hedayat Abbastabar, Foad Abd-Allah, Ahmed Abdelalim, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *The Lancet*, 396(10258):1204–1222, 2020.

Stefan Wager and Kuang Xu. Experimenting in equilibrium. *Management Science*, 67(11):6694–6715, 2021.

G.G. Wagner, R.V. Burkhauser, and F. Behringer. The English language public use file of the German socio-economic panel. *Journal of Human Resources*, 28:429–433, 1993.

Jon Wakefield. *Bayesian and frequentist regression methods*. Springer Science & Business Media, 2013.

Larry Wasserman. *All of statistics: a concise course in statistical inference*, volume 26. Springer, 2004.

Rainer Winkelmann. Markov chain monte carlo analysis of underreported count data with an application to worker absenteeism. *Empirical Economics*, 21(4):575–587, 1996.

Rainer Winkelmann. *Econometric analysis of count data*. Springer Science & Business Media, 2008.

Rainer Winkelmann and Klaus F Zimmermann. *Poisson-logistic regression*. Volkswirtschaftl. Fakultät d. Ludwig-Maximilians-Univ. München, 1993.

Jonathan S Wood, Eric T Donnell, and Christopher J Fariss. A method to account for and estimate underreporting in crash frequency research. *Accident Analysis & Prevention*, 95:57–66, 2016.

Kevin Wood, Satoshi Nishida, Eduardo D Sontag, and Philippe Cluzel. Mechanism-independent method for predicting response to multidrug combinations in bacteria. *Proceedings of the National Academy of Sciences*, 109(30):12254–12259, 2012.

Kevin B Wood. Pairwise interactions and the battle against combinatorics in multidrug therapies. *Proceedings of the National Academy of Sciences*, 113(37):10231–10233, 2016.

Kaan Yilancioglu and Murat Cokol. Design of high-order antibiotic combinations against m. tuberculosis by ranking and exclusion. *Scientific reports*, 9(1):1–11, 2019.

Corwin M Zigler and Georgia Papadogeorgou. Bipartite causal inference with interference. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 36(1):109, 2021.

Anat Zimmer, Itay Katzir, Erez Dekel, Avraham E Mayo, and Uri Alon. Prediction of multidimensional drug dose responses based on measurements of drug pairs. *Proceedings of the National Academy of Sciences*, 113(37):10442–10447, 2016.