Robust Control of Evolutionary Dynamics

Thesis by

Vanessa D. Jonsson

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2016

(Defended July 10, 2015)

© 2016
Vanessa D. Jonsson
All Rights Reserved

For my mother, Ana Maria.

Acknowledgments

I am humbled by the support of my main advisor Richard Murray, who was the principal reason for my success through the relatively daunting yet extraordinary experience of my graduate research. He provided an incredible amount of intellectual freedom and support to pursue my own ideas and research projects, despite their tangential association to his own objectives. His intellectual enthusiasm, integrity and ability to excel at the highest levels in multiple areas of engineering and science has been the inspiration to push the limits of my own knowledge while having strong personal principles as a researcher. I am indebted to my second advisor, David Baltimore, for giving me the immense opportunity to work in his lab five years ago, despite my lack of experience in experimental biology, and in particular with live HIV. His belief in my capabilities, support and enthusiasm for my research project catapulted my understanding of the power and limitations of engineering techniques in the application to biomedical problems. This total immersion among experts in virology and immunology was an incredibly unique and humbling experience. I would like thank Pamela Bjorkman, whose invaluable advice and support was critical to the success of several of my research projects. Her open-mindedness and willingness to support my research ideas was a constant source of encouragement and validation. I am indebted to her for providing such a welcoming environment. I also wish to thank John Doyle for his constant support throughout my graduate studies and beyond and for continually challenging and supporting me in the most difficult of times.

One of the professors that was most influential in the earlier years of my graduate research was Rob Phillips, who opened up a new world of physical biology for me. This eventually led me to connect concepts from equilibrium thermodynamics to the

dynamics of evolution. I am grateful for his commitment to excellent teaching. I am also thankful to Christopher Snow and Justin Bois, who provided useful guidance in the areas of computational chemistry and equilibrium thermodynamics. I am indebted to my collaborators Anders Rantzer and Nikolai Matni, whose help was invaluable in the development of the control theoretic algorithms in this thesis. I would like to thank Anthony West who served as an informal advisor and whose knowledge of the structural biology of HIV antibodies was central in my progress and understanding.

I am thankful for all the support of my former colleagues in both the Baltimore and Bjorkman labs who provided immense technical support and mentoring. I appreciate the help of Alex Sigal who initially took me under his wing and showed me how to navigate HIV tissue culture. I am indebted to Collin Kieffer, who provided technical guidance and help with advanced HIV protocols and to Michael Bethune for his help and advice on experimental design and molecular biology protocols. I am thankful for the helpful discussions with Ron Diskin, Louise Scharf, Jennifer Keeffe, and the support from Geoffrey Lovely, Rachel Galimidi, Jocelyn Kim and Stella Ouyang. I am grateful for my colleagues in the Murray group who have provided useful insight and been supportive throughout the years. I thank Mark E. Davis at Caltech, for his supportive mentoring and encouragement to avidly pursue my research ideas.

Finally, this journey would not have been possible without the endless support of my wonderful family, my mother Ana Maria and father Russell, my stepmother Patricia and my brother Russell. Above all, I want to thank my partner and best friend Nikolai for his unwavering support – every moment of this journey was made better with you by my side.

Contents

Abstract

The application of principles from evolutionary biology has long been used to gain new insights into the progression and clinical control of both infectious diseases and neoplasms. This iterative evolutionary process consists of expansion, diversification and selection within an adaptive landscape — species are subject to random genetic or epigenetic alterations that result in variations; genetic information is inherited through asexual reproduction and strong selective pressures such as therapeutic intervention can lead to the adaptation and expansion of resistant variants. These principles lie at the center of modern evolutionary synthesis and constitute the primary reasons for the development of resistance and therapeutic failure, but also provide a framework that allows for more effective control.

A model system for studying the evolution of resistance and control of therapeutic failure is the treatment of chronic HIV-1 infection by broadly neutralizing antibody (bNAb) therapy. A relatively recent discovery is that a minority of HIV-infected individuals can produce broadly neutralizing antibodies, that is, antibodies that inhibit infection by many strains of HIV. Passive transfer of human antibodies for the prevention and treatment of HIV-1 infection is increasingly being considered as an alternative to a conventional vaccine. However, recent evolution studies have uncovered that antibody treatment can exert selective pressure on virus that results in the rapid evolution of resistance. In certain cases, complete resistance to an antibody is conferred with a single amino acid substitution on the viral envelope of HIV.

The challenges in uncovering resistance mechanisms and designing effective combination strategies to control evolutionary processes and prevent therapeutic failure apply more broadly. We are motivated by two questions:

- Can we predict the evolution to resistance by characterizing genetic alterations that contribute to modified phenotypic fitness?

- Given an evolutionary landscape and a set of candidate therapies, can we computationally synthesize treatment strategies that control evolution to resistance?

To address the first question, we propose a mathematical framework to reason about evolutionary dynamics of HIV from computationally derived Gibbs energy fitness landscapes — expanding the theoretical concept of an evolutionary landscape originally conceived by Sewall Wright to a computable, quantifiable, multidimensional, structurally defined fitness surface upon which to study complex HIV evolutionary outcomes.

To design combination treatment strategies that control evolution to resistance, we propose a methodology that solves for optimal combinations and concentrations of candidate therapies, and allows for the ability to quantifiably explore tradeoffs in treatment design, such as limiting the number of candidate therapies in the combination, dosage constraints and robustness to error. Our algorithm is based on the application of recent results in optimal control to an HIV evolutionary dynamics model and is constructed from experimentally derived antibody resistant phenotypes and their single antibody pharmacodynamics. This method represents a first step towards integrating principled engineering techniques with an experimentally based mathematical model in the rational design of combination treatment strategies and offers predictive understanding of the effects of combination therapies of evolutionary dynamics and resistance of HIV. Preliminary *in vitro* studies suggest that the combination antibody therapies predicted by our algorithm can neutralize heterogeneous viral populations despite containing resistant mutations.

Chapter 1

Introduction

"The power of Selection, whether exercised by man or brought into play under nature through the struggle for existence and the consequent survival of the fittest, absolutely depends on the variability of organic beings. Without variability, nothing can be effected; slight individual differences, however, suffice for the work, and are probably the chief or sole means in the production of new species." - Charles Darwin, 1868.

1.1 Background and Motivation

The modern synthesis. The core of current evolutionary theory was forged seventy years after Charles Darwin's *On the Origin of Species* [20], when statisticians and geneticists began laying the foundation for what is now called the 'modern synthesis'. This allowed the process of evolution to be described mathematically as the change in frequencies of genetic traits in a population over time, uniting Darwin's concept of natural selection with a newly formed field of Mendelian genetics [30, 33, 103]. This theoretical foundation and its corresponding quantitative methods provided support to better understand the tenets of evolutionary theory — that variation arises through random genetic mutation, is inherited by offspring and these together with natural selection leads to adaptation and speciation.

Evolutionary biologists have since expanded upon the modern synthesis framework, drawing concepts and methods from other fields. The discovery of DNA as the material foundation for the encoding and the hereditary transmission of genetic

information, became the driving force of evolutionary theory. It transformed the notion that selection was associated with phenotypic trait variation to being a function of variations driven by genetic mutations. Since then, genetic mapping for numerous phenotypic adaptations across multiple organisms have been established — such as the evolution of influenza virus resistance to antivirals [80], the adaptation of wing shape in [74] and the genetic basis resistance of malaria in humans [34, 39].

Evolution and disease. More recently, the application of evolutionary concepts has been extended to the study of both infectious and neoplastic disease, giving new insights into their progression and clinical control [62, 71, 72]. The underlying processes are equivalent to other evolutionary models — cells or virus are subject to random genetic mutation that result in variations; genetic information is inherited through asexual reproduction and strong selective pressures such as therapeutic intervention can lead to the adaptation and expansion of resistant variants. Genetic instability of neoplasms drives single base sequence changes, chromosomal rearrangements and gene fusion, and these can confer a selective advantage when the resulting phenotype exhibits both a proliferative advantage and a defect in DNA repair [55]. Non-genetic adaptation and phenotypic plasticity can be induced via oncogene inhibition, enabling the survival in cancer cells during initial therapy and thereby promoting residual disease. Recent studies in EGFR mutant lung adenocarcinoma reveal that NF-κB signaling is rapidly engaged upon initial EGFR inhibitor treatment to promote tumor cell survival [7].

The 1987 FDA approval of azidothymidine (AZT), a nucleoside analog reverse-transcriptase inhibitor (NRTI) for the treatment of chronic HIV infection, was one of the first signs of therapeutic promise in the treatment of chronic human immunodeficiency virus type 1 (HIV-1) infection — this treatment significantly reduced viral replication in patients and led to clinical improvements [29]. However, the ability of HIV to rapidly evolve drug resistance was soon observed in patients treated with AZT— the genetic basis for their resistance was explained with the existence of three amino acid substitutions in the reverse transcription gene [50]. It was not until the 1995-1996, that that the subject of resistance was addressed with the development of

a new class of protease inhibitors (PIs) [61], that when combined with two NRTIs, introduced the concept of highly active antiretroviral therapy (HAART). The treatment of chronic HIV infection with HAART is considered one of the great successes of modern medicine in that it radically changed clinical outcomes successfully reducing both patient viral loads to virtually “undetectable" levels, transforming HIV from a fatal disease to one that is a manageable chronic illness [18].

The success of antiretroviral combination therapy in controlling evolution of HIV has provided insight into how the evolutionary processes of other disease models could be controlled for progression and management of therapeutic resistance [4, 94, 99]. With the recent introduction of targeted therapies for the treatment of certain cancers [36], new questions surrounding the effectiveness of tailoring treatments to an individual patient's tumor and its implications with respect to the emergence of driver mutations and resistant phenotypes are being raised. Specifically, these small molecule inhibitors and monoclonal antibodies exploit particular genetic addictions and vulnerabilities of cancer cells, establishing an environment in which the occurrence of mildly drug resistant cells can develop an evolutionary advantage over those for which the therapy is targeted [23, 27, 41, 97]. Clonal expansion of these evolutionary advantageous cells is exacerbated by the presence of considerable genetic intra-tumor heterogeneity already present in treatment-naive patients, contributing to resistance and the need for principled approaches to the design of combination targeted treatment strategies.

Toward a Principled Design of Treatment Strategies. These disease models illustrate more generally the challenges in uncovering resistance mechanisms and designing effective combination strategies to control evolutionary processes that lead to resistance. Similar evolutionary processes are involved in the context of HIV-1 and selection of resistant mutants with respect to broadly neutralizing antibody (bNAb) therapy. In this thesis, we focus our attention to this particular application, but the mathematical techniques that we propose are relevant to other infectious and non-infectious diseases where growth, mutation and selection are central.

Recent evolution studies on HIV-1 have uncovered that bNAb monotherapy can

exert selective pressure on virus resulting in the rapid appearance and evolution of resistant mutant viral population [22, 47, 100]. In certain cases, complete resistance to a bNAb is conferred with a single amino acid substitution on the viral envelope (Env). To address HIV evolution in this context, combination bNAb therapy has been proposed and shown to effectively control infection and suppress viral loads below detection in murine models [22, 47, 100].

Advances in the identification and engineering of anti-HIV-1 antibodies have produced a large set of detailed molecular structures and neutralization data generated against a broad panel of HIV-1 strains. Can we computationally predict evolution to resistance by characterizing genetic alterations that contribute to modified phenotypic fitness? To address this question, we propose a computational model to reason about evolutionary dynamics of HIV from computationally derived fitness landscapes—linking the notion of genotype to phenotype in a quantifiable manner.

The second question we ask is, given a fitness landscape and a set of candidate therapies, can we computationally synthesize treatment strategies and control evolution? The rational design of combination antibody therapies for HIV treatment involves the exploration of a large mutational space in the context of an ever growing number of candidate antibodies — experimentally screening their combinations and concentrations to effectively control evolution to resistance becomes increasingly infeasible. To address this, we require a scalable methodology that can take into account increasing amounts of HIV/bNAb resistance data, bNAb pharmacodynamic models and HIV mutational dynamics. To this end we propose a scalable and computationally tractable algorithm that solves for optimal combinations and concentrations of bNAbs to neutralize virus in light of viral evolution while simultaneously allowing the designer to tailor treatment strategies in light of viral composition, maximum achievable doses, number of bNAbs used and ability to support pharmacodynamics/pharmacokinetic fluctuations, modeling and experimental error.

We briefly discuss prior computational work in addressing this problem, and follow with the thesis outline and contributions.

1.2 Prior Work

Mathematical approaches. There have been numerous attempts to address evolution of resistance with respect to therapeutic intervention from a mathematical perspective within multiple disease contexts. Many computational results address evolution to resistance by employing analytic methods and/or simulations on small scale stochastic evolutionary dynamics models. The Michor lab [63] recently showed the effects of different *erlotinib* dosing strategies in the presence of pharmacokinetic fluctuations on the evolution of resistance of non small cell lung cancer through simulations of a stochastic evolutionary dynamics model. To address tumor heterogeneity by rational combination therapy design, Zhao et al. [105], propose a static multi-objective optimization formulation that is agnostic to evolutionary dynamics but that models the effectiveness of independently acting, additive drugs on different initial tumor populations. Proposed combination treatments were confirmed experimentally for different tumor initial conditions in a murine lymphoma model [106].

The first models describing evolutionary processes in the context of HIV, were inspired largely in part by Manfred Eigen's original quasi species model [26] and have since been proposed to study evolution, antigenic drift [70, 71]. Recent results by Rosenbloom et al. [87], show that simulations of an evolutionary dynamics model of HIV infection subject to changes in antiretroviral dynamics due to adherence are consistent with clinical studies.

Control theoretic approaches. The challenge of designing treatment protocols that prevent escape is one that has been addressed by control theoretic methods. For cancer therapy, results in this spirit apply methods from optimal and receding horizon control [3, 16], as well as gain scheduling [2], to synthesize treatment protocols that are robust to parameter uncertainty, an inherent issue in all biological systems. In the context of HIV and antiretroviral therapy, Hernandez-Vargas et al. [25] propose a discrete time formulation that allows for the design of switching therapy strategies to delay the emergence of highly resistant mutant viruses. There have been several attempts as well to address nonlinear HIV infection dynamics using model predic-

tive control (MPC) to design optimal antiretroviral drug dosing strategies [56, 108]. Recent results in [84] and [8] consider a simplified bilinear model and the optimal control problem is shown to be convex over a finite horizon for a predefined set of initial states.

1.3 Thesis Contribution and Outline

In this thesis, we explore two questions in the evolutionary biology of disease: Can we predict the evolution to resistance by characterizing genetic alterations that contribute to modified phenotypic fitness? Given a fitness landscape and a set of candidate therapies, can we computationally synthesize treatment strategies and control evolution? We focus our application to the antibody treatment of chronic HIV infection, but the mathematical techniques that we propose are relevant to other infectious and non-infectious diseases. Many of the contributions of the following chapters are based on a number of publications [42, 43, 44], indicated below.

Chapter 2: Evolutionary Dynamics on Computationally Derived Fitness Landscapes. We propose a computational model to reason about evolutionary dynamics of HIV on computationally derived fitness landscapes. Our approach combines well-utilized HIV dynamical systems models, incorporates infection and antibody neutralization dynamics, a mutation process, and a method that uses energy minimization calculations on structural information to quantify fitness differences between sensitive and resistant strains. Specifically:

1. We propose and develop an extension of the least absolute shrinkage and selection operator (LASSO) to identify mutational phenotypes and uncover potential escape mutants from neutralizing anti HIV antibodies.

2. We develop a biophysical model based on Gibbs free energy of binding derived from energy minimization calculations on structural information to quantify fitness differences between sensitive and resistant HIV strains.

3. We develop an HIV evolutionary dynamics model to include infection and neutralization reaction rates based on computed Gibbs energy fitness landscapes.

Chapter 3: Robust Control of Evolutionary Dynamics. Chapter 3 presents three algorithms for the principled design of targeted combination drug treatment strategies that explicitly account for the evolutionary dynamics of a generic disease model, where the drugs under consideration are non-interacting and exhibit independent additive effects. These algorithms allow the designer to quantifiably explore tradeoffs between number of therapies used (controller sparsity), therapy concentrations (magnitude of the gain) and ability to support pharmacokinetic fluctuations (robustness to perturbations). Our contribution specifically is itemized below.

1. Our first algorithm proposes a general iterative method that uses an $\mathcal{H}_\infty$ robust control approach to design targeted combination therapy concentrations and is effective in generating robustly stabilizing controllers.

2. Our second algorithm addresses large scale systems concerns lacking in the first algorithm, presenting a scalable solution to the combination therapy problem by reformulating it as a second order cone program (SOCP), with robustness guarantees addressed by minimization of the induced $\mathcal{L}_1$ norm.

3. Our third algorithm solves the combination therapy problem subject to the same design constraints (sparsity of the drug combination, maximum dosage and robustness constraints) formulated as an SOCP while addressing the nonlinear dynamics of individual drugs and of their combinations.

Chapter 4: Engineering Antibody Treatment Strategies to Control HIV. We demonstrate our ability to control the evolution to resistance of HIV in the presence of antibody therapy, through the application of the combination therapy algorithms developed in Chapter 3 as applied to experimental data derived from recent published studies [14, 22, 47]. We also discuss a preliminary *in vitro* experimental methodology and results and show that the antibody treatment strategies synthesized with the nonlinear pharmacodynamics combination therapy algorithm described in

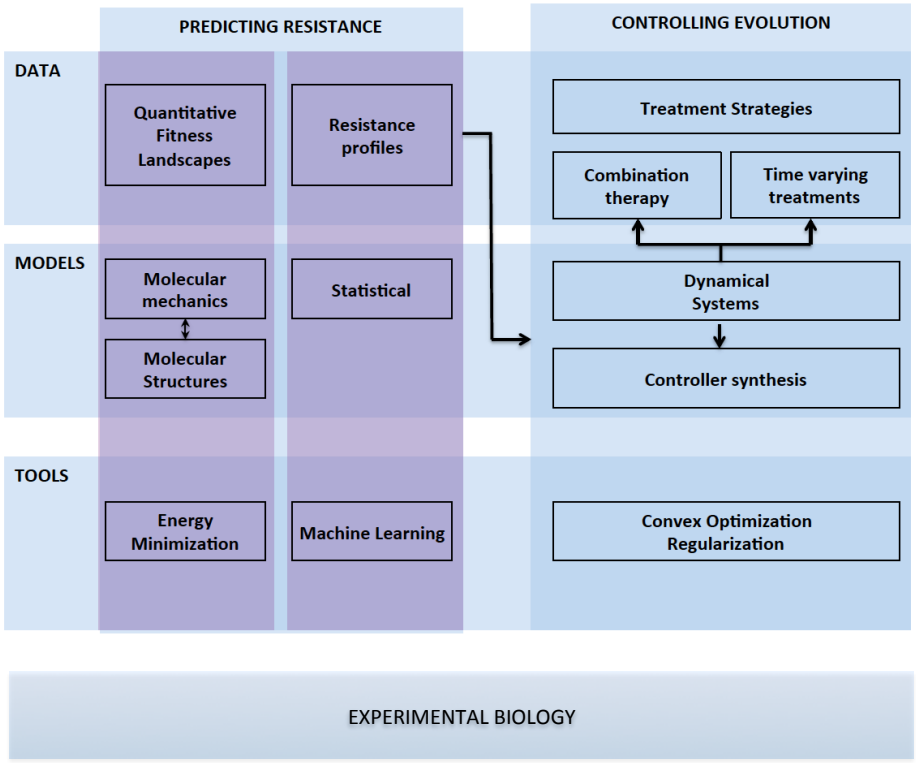

Figure 1.1: Overview of the data, mathematical models and tools used in this thesis.

Section 3.5.3 controls infection despite the presence of a mixed initial population of viruses, most of which are resistant to at least one antibody in the mix. Specifically:

1. We synthesize combination treatments and compare the respective $\mathcal{H}_\infty$ and the $\mathcal{L}_1$ combination therapy algorithms with respect to their performance and robustness to biologically relevant uncertainty models and unmodeled dynamics.

2. We develop a high throughput *in vitro* experimental system to identify the replication and neutralization properties of HIV mutants and populate parameters for our dynamical systems model, as well as test our predicted bNAb combination therapies.

3. We demonstrate successful *in vitro* validation of our computationally predicted bNAb combinations on heterogeneous viral populations comprised of resistant mutants.

Chapter 2

Evolutionary Dynamics on Computationally Derived Fitness Landscapes

2.1 Introduction

In evolutionary biology, the concept of fitness landscape serves to associate genotype to some measure of fitness, or phenotype. With the growing number of detailed molecular structures and recent advances in modeling and computational approaches, genotype-phenotype relationships are now being quantified and generated in an automated way. Thus, fitness landscapes are transitioning from a concept used for visualization of fitness distributions, to computable, quantifiable, multidimensional fitness surfaces upon which to study complex evolutionary outcomes.

Recent pre-clinical and clinical studies recently demonstrated that HIV can escape from antibody mono therapy, and in some cases, combination therapy [14, 37, 47, 95]. In this chapter, we develop a computational framework that explains these observations and predicts the likelihood of certain resistant mutants in the presence of antibody mono therapy. Our method uses energy minimization calculations on structural information and statistical inference on antibody neutralization data to quantify fitness differences between sensitive and resistant strains and incorporates this data into an HIV dynamical systems model of infection, mutation and antibody neutralization. We show that our evolutionary dynamics simulations on computationally generated

HIV fitness landscapes have reasonable agreement with experimental findings. This represents a first step in modeling and predicting HIV escape from antibody therapy but has a broader application in evolutionary dynamics settings where a quantitative relationship between genotype and evolutionary fitness can be established.

The chapter is organized as follows: in Section 2.2.1, we propose a statistical inference method to uncover correlations between HIV-1 viral envelope (Env) sequence data and antibody neutralization data. In Section 2.2.2, we compute HIV-1 fitness landscapes using energy minimization techniques and argue that quantifying changes to both infection and neutralization due to mutations are equally important aspects in the study of HIV evolution to antibody resistance. In Sections 2.2.2 and 2.2.3, we derive generalized Hill equations to express bound gp160/CD4 and gp160/antibody as a function of the differences in Gibbs free energy of binding due to point mutations. Section 2.2.4 connects preceding work by incorporating computed fitness landscapes into a stochastic HIV evolutionary dynamical systems model of infection, mutation and antibody neutralization and illustrates the applicability of these methods for the prediction of resistance dynamics in light of antibody monotherapy for HIV.

2.2 Results

2.2.1 Statistical inference to uncover resistance phenotypes

We developed a statistical model and used this to uncover correlations between HIV-1 envelope (Env) sequence data and antibody neutralization data. In order to obtain a model that could explain antibody neutralization data with a minimal set of residues while taking into account irregularity in the experimental data set corresponding to the limit of neutralization assays (Figure 2.1), we extended the well known least absolute shrinkage and selection operator (Lasso) [98]. Our model, the saturated Lasso (satlasso), returns a set of amino acid residues and their Env sequence location that explains antibody neutralization data by minimizing error between model and experimental data, penalizing large models, and appropriately weighing errors

corresponding to saturated experimental data. A mathematical description of the model can be found in Section 2.4.1. Model selection is performed with 5-fold cross validation [79].

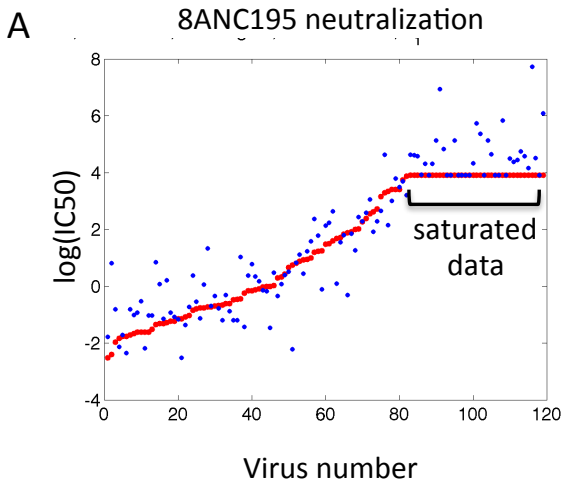

Figure 2.1: **A.** An example of the satlasso estimator as applied to 8ANC195 antibody neutralization data. Saturated data is due to the limit of the neutralization assay and is modeled in Equation (2.6). Red points correspond to experimental data, blue points correspond to the estimated model.

To assess the generalization of our satlasso model, we compare the first six largest magnitude regressors to experimentally derived data sets and find that many Env residues known to be critical to neutralization are identified (Figure 2.2). For the gp120-gp41 bridge antibody 8ANC195, changes in glycosylation at 276, 234 and 230 sites are found to induce large changes in neutralization [91]. Specifically, the introduction of a glycan at position 230 in a wild-type YU2 HIV strain, leads to a sixfold increase in IC50, whereas removal of 234 and 276 sites leads to a more substantial increase in IC50 [91, 101]. Our model identifies this epitope and captures the corresponding changes in IC50 through the magnitude of each regressor 276++ (-2.79), 234++ (-2.77) and 230++ (1.59) (Figure 2.2).

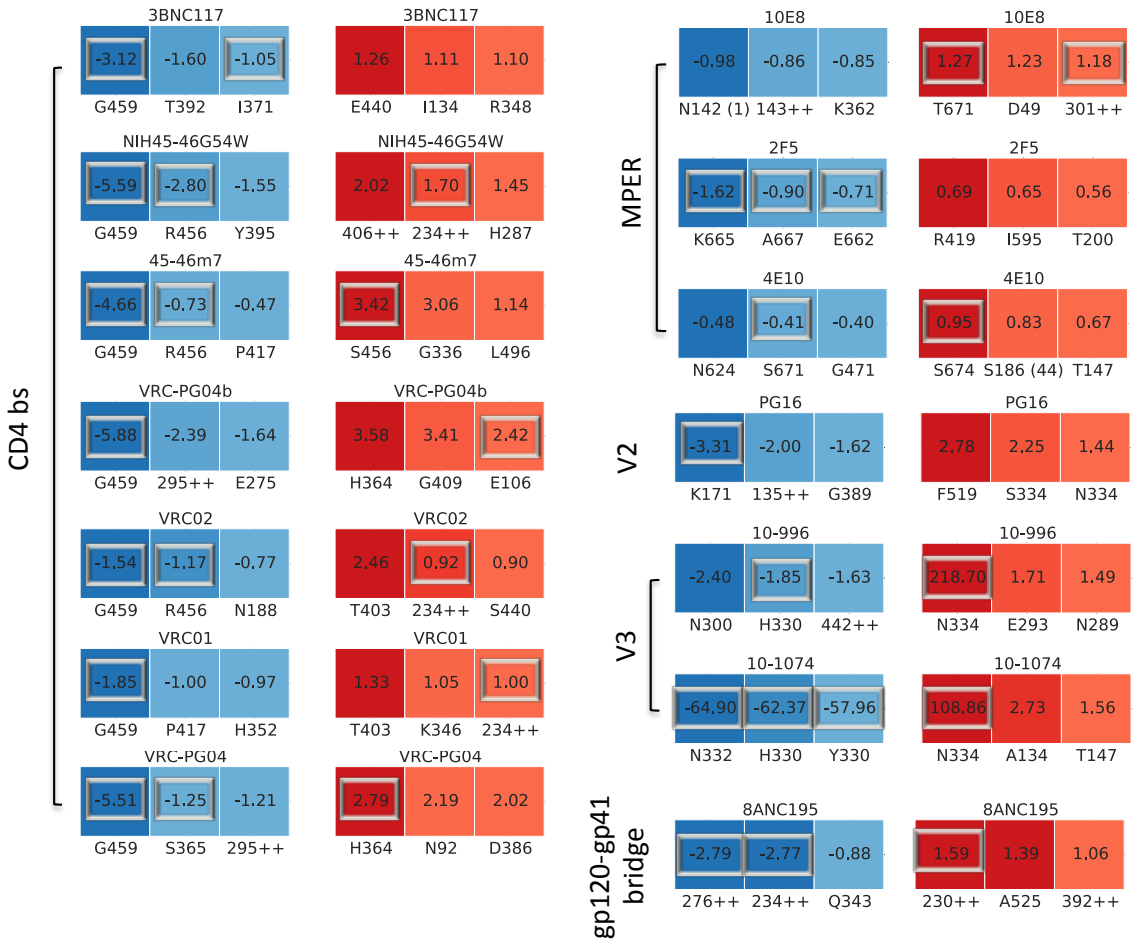

Figure 2.2: The first six regressors with largest magnitude contributing to neutralization (blue) and resistance (red) for each of CD4 binding site, V2, V3, MPER and gp120-gp41 bridging antibodies in this study, solved with satlasso using 5-fold cross validation. Overlaid boxes correspond to amino acid changes that have been validated experimentally.

For all CD4 binding site antibodies in this study, we find that the existence of a Gly at position 459 is most important for neutralization, confirming previous evolution studies that mutations at position 459 confer resistance in 3BNC117 and NIH-4546-G54W [47, 22]. Our model uncovers that an Arg at position 456 is also significant for neutralization but less so than the site at position 459. Recently, Lynch et al. show that a Trp mutation at position 456 has a modest effect on neutralization by many CD4-bs antibodies with the exception of VRC-PG20 [57].

Functional studies on V3 loop binding antibodies 10-996 and 10-1074 reveal an essential dependence on the N332 associated glycan [66]. Our model identifies this

critical site and shows that the existence of a His or Tyr at position 330 is necessary for neutralization of 10-1074. There is some evidence that the H330Y substitution has no effect on neutralization of either PGT121 and 10-1074 when this mutation was tested on Simian-Human Immunodeficiency Virus, SHIVAD8 [89].

The highly conserved membrane proximal external region (MPER) of the gp41 transmembrane subunit on HIV-1 has recently been linked to epitopes of several broadly neutralizing antibodies 2F5, 10E8 and 4E10 [38, 88, 65]. Alanine scanning, structural and paratope analysis each indicate that 10E8 makes crucial contacts with highly conserved residues W672, F673, W676 and K/R683 [38]. Although our model does not recover these particular contact sites, it does identify a Thr substitution at 671 that is shown to raise IC80 values above 20 μg/ml in otherwise sensitive JR2 virus [38]. For 4E10, our model captures both one epitope site at 671 and a known resistant substitution at 674S [9]. The 2F5 epitope is defined by a linear segment of gp41 residues 662 - 668 with the key binding residues at N664, K665 and Y666 [6, 109]. Our model recovers the crucial binding site at position 665 and two sites on the epitope, E662 and A667 that each confer resistance at the IC90 level with Ala and Gly substitutions respectively [12, 109].

Our results demonstrate that satlasso recovers residues that are critical for neutralization and those that contribute to resistance for a large class of anti-HIV-1 antibodies. Despite the fact that substitutions in certain positions can mediate conformational or other effects within Env that may be best represented by a nonlinear model, some changes in neutralization can nonetheless be captured by this linear regression model.

Recent studies in viral fitness costs associated with escape mutants from the class of CD4-bs antibodies show that viral replicative fitness may be diminished with certain single mutations on the CD4 receptor binding site on gp120 [57], Chapter 4. This suggests more broadly that evolution to resistance can be viewed as a function of both viral replicative and antibody neutralization fitness. We elaborate this idea in the following section and develop a method that quantifies both aspects of viral fitness and utilizes it to study the evolutionary dynamics of HIV in the presence of

antibody monotherapy.

2.2.2 Gibbs Energy Landscapes Correlate With Known Escape Mutations

We hypothesize that a virus' capacity to infect and be neutralized by specific neutralizing antibodies (NAbs) can be approximated by differences in Gibbs free energy of binding associated to *de novo* point mutations on the envelope glycoprotein (Env) complexed with the CD4 receptor and antibody structure. To compute fitness landscapes relating to viral replication and antibody neutralization, we apply an empirical force field, Fold X [93], to evaluate the effect of point mutations on the stability, folding and dynamics on detailed Env/CD4 and Env/NAb molecular structures.

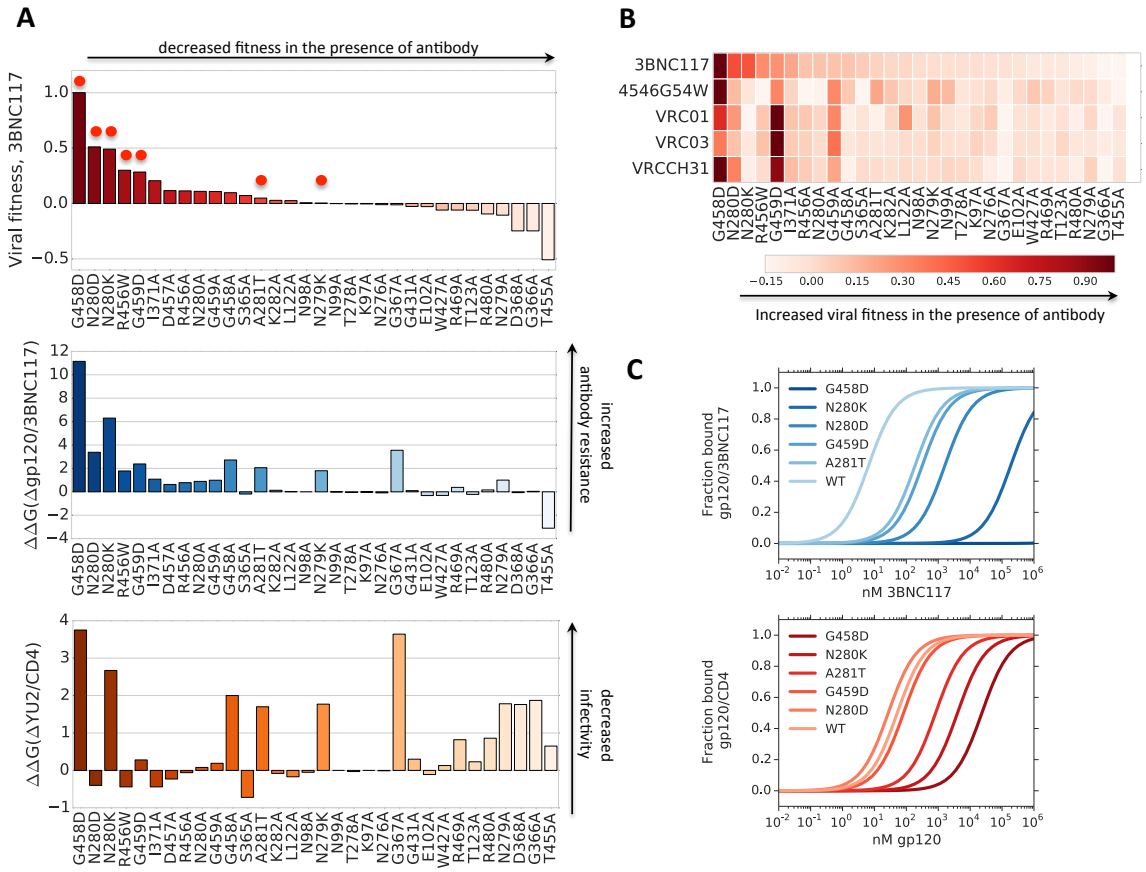

Figure 2.3: **A**. Gibbs binding energy differences computed between wild type and a subset of gp120 point mutations consisting of alanine substitutions on a subset of 3BCN117 antibody resistant mutations. Viral fitness $\mathcal{F}_{\Delta \mathrm{gp}/NAb}$, is computed by Gibbs energy of binding differences between neutralization and infection reactions and normalized with respect to largest value (Equation 2.3) (above). Gibbs energy differences for Δgp120/3BCN117 binding reaction (middle) and the Δgp120/CD4 interaction (below). All mutations are numbered with respect to the HXBC2 reference genome. **B**. Viral fitness computed as in **A** for CD4 binding site antibodies 3BNC117, 4546G54W, VRC01, VRC03, and VRCCH31. **C.** Fraction bound gp120/3BNC117 and CD4/gp120 as computed with Equations (2.5) using Gibbs energy landscapes for antibody and CD4 binding.

The resulting Gibbs free energy landscape provides an equilibrium thermodynamic representation of the fitness of each point mutant with respect to the following simplified binding reactions for infection and neutralization:

$$\mathrm{gp} + \mathrm{cd4} \longrightarrow \mathrm{gp} \cdot \mathrm{cd4}, \tag{2.1}$$

$$\ell + \mathrm{gp} \longrightarrow \ell \cdot \mathrm{gp}, \tag{2.2}$$

where the HIV membrane glycoprotein gp, binds to the CD4 receptor, denoted cd4, in reaction (2.1), and to antibody ℓ during neutralization in reaction (2.2).

We compute the difference in Gibbs free energy of binding between mutant and wild type viral glycoprotein gp160,

$$\Delta\Delta\mathcal{G} = \Delta G_{mut} - \Delta G_{wt},$$

for all point mutations on the solved portions of gp160 in both CD4 and antibody complexes listed in Table 2.2, Section 2.5.

To quantify the effects of mutations on both infection and neutralization, we define viral fitness $\mathcal{F}_{\Delta\text{gp}/NAb}$ as a function of the cost of the mutation (Δgp) with respect to antibody binding minus its cost with respect to CD4 binding

$$\mathcal{F}_{\Delta\text{gp}/NAb} = \Delta\Delta G(\Delta\text{gp/NAb}) - \Delta\Delta G(\Delta\text{gp/CD4}). \tag{2.3}$$

To illustrate this measure of viral fitness, we compute binding energy landscapes for CD4 binding site (CD4bs) antibody 3BNC117 using gp120/3BNC117 (PDB ID: 4JPV) and gp120/CD4 (PDB ID: 1G9N) structures and find reasonable agreement between this measure and known resistant mutations in the presence of 3BNC117 [37, 57] (Figure 2.3, A). In particular, we note that the presence of either an Asp at position 458 or an Ala at position 367 decreases binding to both 3BNC117 and to CD4, simultaneously affecting the virus' ability to infect and to be neutralized by antibody. Binding energy differences between mutations and the CD4 receptor are similar for both G367A and G458D (3.64 kcal/mol and 3.75 kcal/mol) whereas they differ significantly for 3BNC117 binding (3.55 kcal/mol (G367A) and 11.15 kcal/mol (G458D)), suggesting that both mutations could exhibit compromised viral replication but that due to the greater difference in 3BNC117 binding, the Asp mutation at location 458 is more likely to escape 3BNC117 neutralization. This is consistent with our viral fitness calculations and model simulations (Figure 2.5, B) as well as previous experimental validation showing that the Asp mutation at position 458 evolves in the presence of 3BNC117 in murine HIV infection models [37, 57].

To extend our analysis to other CD4bs antibodies, we compute Gibbs binding energy landscapes for 3BNC117, 4546G54W, VRC01, VRC03 and VRCCH31. Our viral fitness calculations uncover that the most resistant substitutions across all CD4bs in the study consist of an Asp substitution at either locations 458, 280 and 459 (Figure 2.3, B). Previous studies show that both the Asp substitutions at location 458 and at location 280 evolve in 4546G54W and 3BNC117 monotherapy experiments [22, 37] and abrogates neutralization by VRCCH31, and VRC01 [57]. The application of our statistical inference model from Section 2.2.1 and previous experimental validation [22, 37, 101] shows the importance of Gly at location 459 for effective neutralization by all CD4bs antibodies.

More subtle differences in viral fitness differences can be uncovered by computing Gibbs energy landscapes. We note that the Trp substitution at location 456 exhibits a more modest increase in viral fitness (0.15-0.45) than the Asp substitution at 458 (0.45-1.0) across the CD4bs antibodies in the study, confirming recent experimental studies that show that the Trp substitution at location 456 modestly decreases but does not completely abrogate neutralization in all CD4bs antibodies [57].

2.2.3 Hill Functions Relate Gibbs Landscapes And Dynamical Systems Parameters

Although analysis of detailed molecular structures may elicit molecular properties by which resistance occurs, the pharmacodynamics of associated drug resistance are less well understood. Therapeutic effect of resistant mutations is typically measured as a change in IC50 relative to wild type. In the case of antiretroviral therapy for HIV, Sampah et al. argue that inhibition can only be predicted if the shape of the dose-response curve is known [90]. This relationship takes the form of a Hill function, an equation extensively used in pharmacology to analyze nonlinear drug-receptor relationships.

We derive generalized Hill equations and express the fraction of bound Env/CD4 and Env/antibody complexes formed during infection and neutralization binding re-

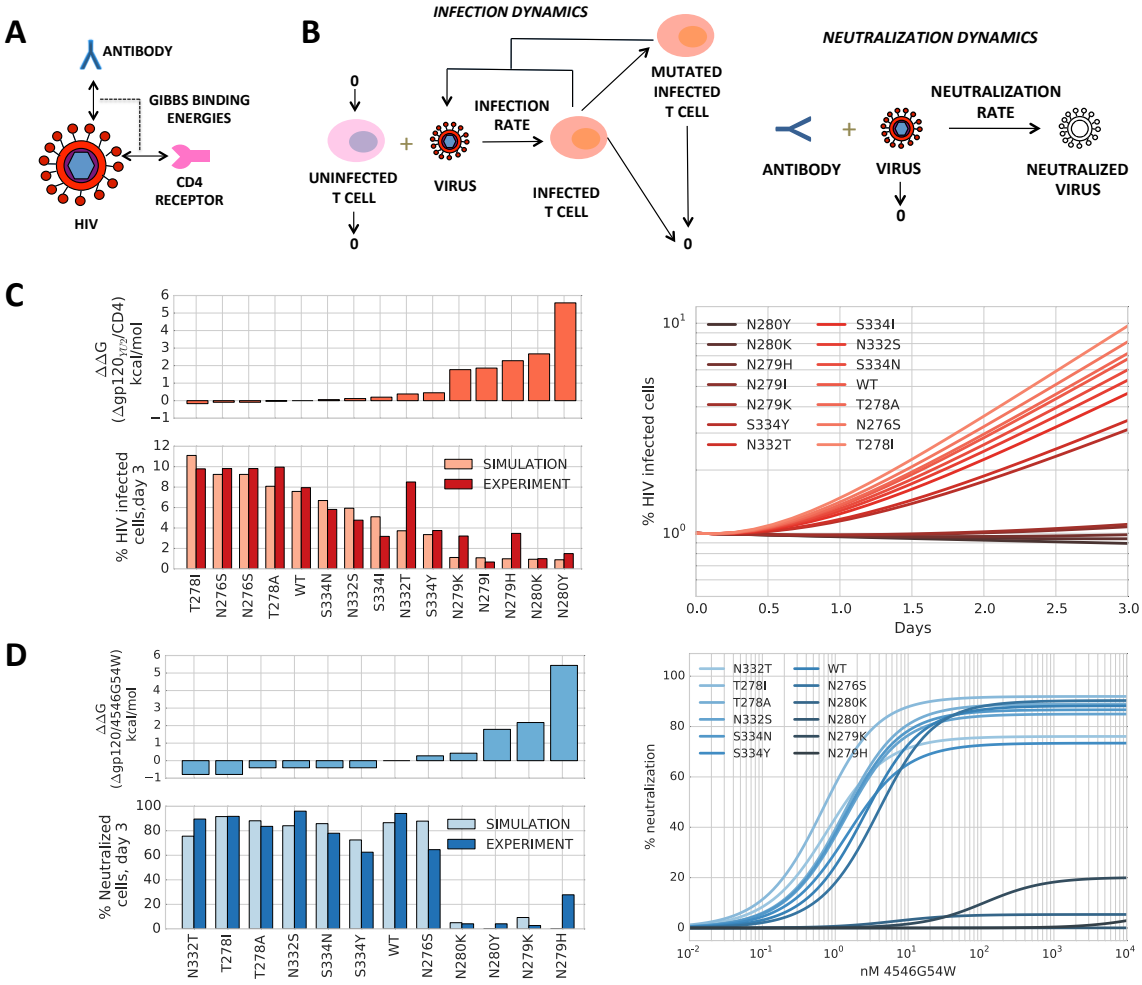

Figure 2.4: **A.** A depiction of the molecular binding model approximating the Gibbs binding energy between viral envelope glycoprotein and the neutralizing antibody and the CD4 receptor, using simplified binding reactions (2.1) and (2.2). **B.** A depiction of the evolutionary dynamics model of HIV infection, mutation and neutralization in the presence of antibody therapy. **C.** (Left) Difference in Gibbs binding energy between wild type and mutations of the viral glycoprotein and the CD4 receptor, for a clade B virus YU2, (PDB ID: 1G9N) for a subset of point mutations. Percent infected cells on day three post secondary infection (dark red) and simulations of the molecular binding plus infection dynamics model using the calculated Gibbs free energy landscape (light red) for a subset of viral glycoprotein mutations. (Right) Three day infection simulations using the calculated Gibbs energies for selected point mutations. **D.** (Left) Difference in Gibbs binding energy between wild type and mutations of the viral glycoprotein and broadly neutralizing antibody 4546G54W, (PDB ID: 4JKP) for a subset of point mutations. Percent neutralized cells on day three in the presence of 10 $\mu g/ml$ 4546G54W (dark blue) and simulations of the molecular binding and infection and neutralization dynamics model using both Gibbs free energy landscapes in the presence of 10 $\mu g/ml$ 4546G54W (light blue) for a subset of viral glycoprotein mutations. (Right) Three day neutralization dynamics simulations using calculated Gibbs energies for selected point mutations for a range of 4546G54W antibody concentrations.

actions, as a function of the differences in Gibbs free energy due to point mutations. We begin by noting that the dissociation constant K_d is an equilibrium value that can be expressed in terms of the free energy of binding, and define ΔK_d^{mut} to be the ratio between mutant and wild type dissociation constants:

$$\Delta K_d^{mut} = \frac{K_d^{mut}}{K_d^{wt}} = \Delta\Delta G/RT, \tag{2.4}$$

where R is the ideal gas constant and T is temperature. The resulting Hill equations model the fraction bound complexes involved in both infection and neutralization reactions:

$$\begin{aligned} R_{mut} &= \frac{[\text{gp}]^m}{[\text{gp}]^m + (K_{inf}^{wt}\Delta K_{inf}^{mut})^m} \\ N_{mut} &= \frac{[\ell]^n}{[\ell]^n + (K_{neut}^{wt}\Delta K_{neut}^{mut})^n}\ , \end{aligned} \tag{2.5}$$

where the symbol [ ] indicates nanomolar concentration. ΔK_{inf}^{mut} and ΔK_{neut}^{mut} are the difference between the dissociation constants for CD4 and antibody binding reactions (2.1) and (2.2) between a mutation *mut* and the wild type viral glycoprotein as defined by Equation (2.4).

Remark 1. We assume that binding associated with the infection and neutralization process is noncooperative, and the corresponding Hill coefficients m and n are approximately 1. Recent experimental studies characterizing antibody neutralization across a diverse virus panel, suggests that for CD4 binding site antibodies $n \approx 0.9 - 1.37$ [92], for V2 antibodies $n \approx 0.7$ and for V3 binding antibodies $n \approx 1.5$. [64].

Figure 2.3 C depicts the Hill equations (2.5) associated with infection and neutralization reactions for a subset of representative 3BNC117 resistant mutations and their computed Gibbs binding energy landscapes. For highly resistant mutations most likely to evolve in the presence of 3BNC117 like the Asp mutation at location 458, we expect a small fraction of bound gp120/3BNC117 regardless of antibody concentration and a moderate concentration of bound gp120/CD4, representing productive infection. More moderately resistant mutations, such as the Thr mutation at 281, requires

two orders of magnitude increase in 3BNC117 concentration to achieve an equivalent fraction of gp120/antibody bound complexes and neutralization as wild type virus. This suggests that at lower concentrations of antibody, viruses with similar fitness profiles as A281T may not be all neutralized by antibody, creating an environment in which this moderately resistant virus can become dominant or acquire additional mutations and achieve greater resistance [22, 37, 47, 57].

2.2.4 Evolutionary Dynamics on Quantifiable HIV-1 Fitness Landscapes

To reason about the dynamics of HIV evolution to resistance in the presence of antibody therapy, we combined computed Gibbs energy landscapes with a stochastic evolutionary dynamical system to model infection, mutation and antibody neutralization (Figure 2.4, A, B). A mathematical description of the virus dynamics and mutation model can be found in Section 2.4.1.

In order to validate of our model, we performed simulations of HIV infection and antibody neutralization and compared these results to replication and neutralization assays performed on gp120 mutations subcloned into a YU2-Env/NL4.3 infectious backbone (Materials and Methods, Section 2.4.3). Both our simulations and experiments were performed under the same conditions and were found to be in agreement, (Figure 2.4, C, D), suggesting that our proposed mathematical model of HIV infection and neutralization based on Gibbs energy landscapes can be used to reasonably predict both infectivity and neutralization rate changes due to point mutations on Env. Simulations of our HIV infection model (Figure 2.4, C) confirm the well-known exponential growth dynamics of HIV infection for mutations with small changes in CD4 binding energies. For mutations with compromised CD4 binding such as the Lys mutation on 279 and the Tyr mutation at 280, our simulations and experiments show significantly decreased infection rates.

To illustrate how different neutralization profiles affect the evolution of resistance, we ran model simulations for a subset of Env point mutations and a range

of 4546G54W concentrations (Figure 2.4, D). We found that certain very resistant mutations are less likely to be affected by high doses of 4546G54W, such as the Lys mutation on 279 and the Tyr mutation at 280, whereas the degree of neutralization of moderately resistant mutations such as the Ser substitution at 279 is dependent on 4546G54W concentration (Figures 2.3, C and 2.4, D). High concentrations of 4546G54W are likely to shift the evolution of viral distributions towards highly resistant mutations by creating an environment in which mutations with moderate resistance are adequately neutralized, leaving the possibility of outgrowth of highly resistant mutants despite low infectivity. We observe this phenomenon in our evolutionary dynamics simulation in the presence of another CD4bs antibody, 3BNC117 (Figure 2.5,B). Specifically, the Lys substitution at location 280 is shown to outgrow all other mutations with lower viral fitness, despite its lower infectivity (Figures 2.4,C and 2.5,B).

To further explore how antibody concentration could influence the composition of viral distributions at steady state, we ran fifty stochastic simulations of our HIV evolutionary dynamics model for different concentrations of 3BNC117 (Figure 2.6,A). We observe that the shape of the stationary distributions varies as a function of antibody concentration with broader peaks occurring at lower antibody concentrations and higher, narrower peaks at high antibody concentrations. We use the Gini coefficient (Materials and Methods, ss:gini) to describe the shape of the stationary viral distribution, and show that it increases as a function of different antibody concentration for multiple CD4bs antibodies (Figure 2.6,B).

To understand which mutants might evolve in the presence of CD4bs antibodies, we ran fifty evolutionary dynamics simulations using computed CD4 and antibody binding energy landscapes for 1664 different mutants of YU2, using constant concentrations of either VRC01, VRC03, 3BNC117 or NIH4546. For a starting population of monoclonal wild type virus, we show that the evolved mutations have close agreement with previously studied escape mutations (Figure 2.5, C). Specifically, our HIV evolutionary dynamics simulations in the presence of NIH-4546 reveal the evolution of mutations at four out of five locations uncovered in recent evolutionary studies

on antibodies NIH-4546 and NIH-4546G54W [47, 22]. Moreover, recent experimental results show that mutations N280D, N280K, A281T, G458D, G459D evolve from 3BNC117 monotherapy [37] - our simulations demonstrate the evolution of four out of these five specific point mutations.

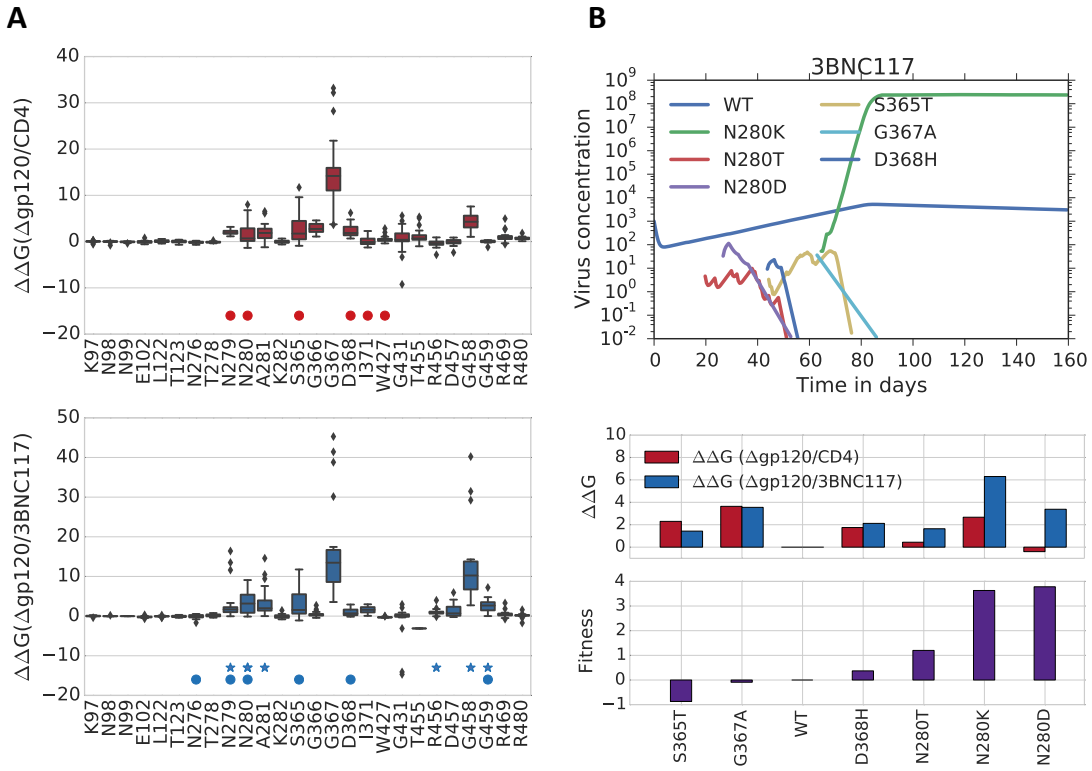

Figure 2.5: **A.** Fitness landscapes representing the difference in Gibbs binding energy between wild type and mutations on the viral glycoprotein (Δgp120) and the CD4 receptor (PDB ID: 1G9N) (above) and the neutralizing antibody 3BNC117 (PDB ID: 4JPV) (below) for all amino acid substitutions on a subset of residues. Closed circles indicate binding site locations, stars indicate residue locations for which resistant mutations have been found. **B.** One simulation of the HIV evolutionary dynamics model (Equation 2.7, Materials and Methods) on computed Gibbs energy landscapes shown in **A** and for a concentration of 5 $\mu g/ml$ 3BNC117. The simulation is run with an initial condition 10^5 uninfected cells/ml, 1000 virions/ml of of wild type YU2 clade B HIV. (Below) Fitness landscape and computed viral fitness for evolved mutations.

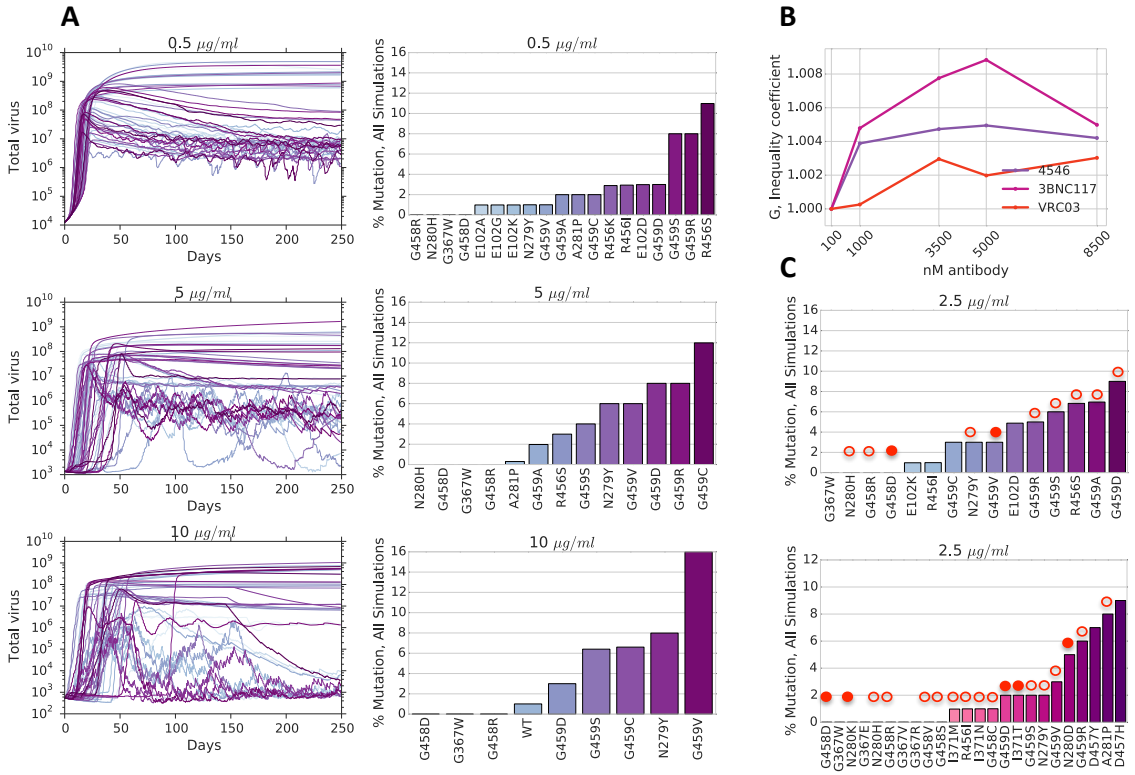

Figure 2.6: **A**. (Left) Fifty simulations of the HIV evolutionary dynamics model (Equation 2.7, Materials and Methods) on computed CD4 and antibody fitness landscapes for different NIH-4546 antibody concentrations. Simulations are run with an initial condition of 10^5 uninfected cells/ml, 1000 virions/ml of of wild type YU2 clade B HIV. (Right) Viral distributions at day 250 for all corresponding simulations, representing the percent mutation present in the total population at steady state. **B**. Gini coefficient for antibodies NIH-4546, 3BNC117 and VRC03 as a function of antibody concentration, for steady state viral distributions in **A**. **C**. Viral distributions at day 250 for all 50 simulations of HIV evolutionary dynamics for (top) 2.5 $\mu g/ml$ of NIH-4546 and (bottom) 3BCN117. Filled circles represent particular mutations that are shown to evolve experimentally [22, 47] or to be resistant [57, 101]. Empty circles represent the residue locations at which mutations either evolve or have compromised neutralization by antibody.

2.3 Discussion

Recent advances in the identification and engineering of anti-HIV-1 antibodies have produced a large set of detailed molecular structures and neutralization data generated against a broad panel of HIV-1 strains. Recent computational analysis of antibody neutralization data has been successful in categorizing antibodies with respect to their neutralization activity [32], extracting the identities of Env residues that are necessary for neutralization [101] and uncovering antibody epitopes [73]. Here, we report a computational methodology that utilizes antibody neutralization data and structural information to construct HIV fitness landscapes and reason about the dynamics of resistance. It consists of the interpretation of neutralization data using statistical inference, the construction of fitness landscapes using computational chemistry and the development of biophysical and mathematical model to capture the dynamics of replication, mutation and selection.

Our statistical model is able to uncover critical Env residues involved in antibody neutralization and is consistent with recent studies in antibody resistance. It does not identify the entire structural epitope involved in the protein-protein contact. Rather, satlasso identifies an epitope that is involved in the function of the protein-protein interaction, in our case neutralization.

To address virus fitness in terms of its replication and neutralization capabilities, we computed Gibbs binding energy landscapes on gp120/CD4 and gp120/NAb structures and found a good correspondence with our experimental studies. One of the drawbacks of using force fields to compute fitness landscapes, is that they rely on experimental data and are therefore empirical. Furthermore, Xray structures are prone to errors due in part to the non-physiological conditions under which the structure is determined. Setting these concerns aside, tools like Fold X are geared specifically toward screening the effect of single nucleotide polymorphisms (SNPs) on protein stability and are therefore well suited for our problem. For non-conservative mutations (those not involving Ala or Gly scanning), a reorganization of the protein backbone is often necessary, requiring an exploration of alternative conformations. However, for

large structure reconfigurations such as a deletion or creation of a glycan structure, we observe that Fold X is not always able to capture the corresponding effects on binding energy. In the case where glycosylation changes have a large effect on neutralization by antibody, satlasso is likely to recover these features, as illustrated by the accurate uncovering of the glycan dependence of 8ANC195. Therefore, the combination of data derived from satlasso and Gibbs energy landscapes offer complementary views on viral fitness.

We present an HIV evolutionary dynamics model that is the first to our knowledge to incorporate binding energy landscapes of replication and neutralization and that accurately predicts the evolution of point mutations in the presence of anti-HIV antibody monotherapy. Recent evolution studies uncovered the evolution of double mutations that conferred resistance to several antibodies [22, 47, 37]. A clear extension to our model is to incorporate double mutations in our Markov model of mutational dynamics however a clear limitation is the accuracy of an empirical energy minimization in the case of such double mutations.

2.4 Materials and Methods

2.4.1 Mathematical Models

satlasso. We define the saturated least absolute shrinkage and selection operator (satlasso) and formulate it as a convex optimization problem. We consider that the antibody neutralization data $X = X_s + X_u$ is comprised of saturated data X_s, corresponding to IC50s of very resistant viruses, and of unsaturated data X_u (Figure 2.1). Observe n predictor response pairs (x_i, y_i) where $x_i \in \mathbb{R}^p$ and $y \in \mathbb{R}$. Forming $X \in \mathbb{R}^{n \times p}$, $X = X_u + X_s$ with standardized columns, the *saturated lasso, (satlasso)* is an estimator defined by the following convex optimization problem:

$$\begin{aligned}&\text{minimize}_{\ \beta\in\mathbb{R}^p}\\&\underbrace{\lambda_1\frac{1}{m}||y_u - X_u\beta||_2}_{\text{Error unsaturated data}} + \underbrace{\lambda_2\frac{1}{m}\sum_{i=1}^{n} w_i|\beta_i|}_{\text{Sparsity}} + \underbrace{\lambda_3\ \max(y_s - X_s\beta, 0)}_{\text{Error saturated data}},\end{aligned} \tag{2.6}$$

where $w_i = \frac{1}{\hat{\beta}_i}$ and $\hat{\beta}_i = \frac{1}{\xi}\sum_{j=1}^{\xi}|\hat{\beta}_i^j|$ where $\hat{\beta}^j$ is the solution to the ordinary least squares problem for subset j.

Model selection was performed by 5 fold-cross validation. Our results exclude mutations that are located in the signal peptide region, transmembrane and intracellular regions corresponding to HXBC2-numbered residues (1-30), (685-706) and (706-end) respectively. These locations on Env are not exposed to antibody binding and we assume that are not subject to selective pressure by antibody.

Viral dynamics. To simulate how HIV might evolve resistance, we extended the widely used HIV infection dynamics model [77] with a random mutation process and included our Gibbs energy formulation to capture the effects of genomic variation on the dynamics on infection and neutralization by antibody therapy. The stochastic discrete time differential equation model of HIV evolution under antibody selection is written as:

$$\begin{aligned}x[k+1] &= \lambda + x[k] - (\eta_c\textstyle\sum_i^n R_i x[k] - d_x x[k])\tau,\\ y_i[k+1] &= y_i[k] + (\eta_c R_i x[k] + \Upsilon_i[k] - d_y y_i[k])\tau,\\ v_i[k+1] &= v_i[k] + ((1-N_i)k_v y_i[k] - N_i v_i[k] - d_v v[k])\tau,\end{aligned} \tag{2.7}$$

where $x \in \mathbb{R}_+$ is the concentration of uninfected CD4$^+$ T cells, $y_i \in \mathbb{R}_+$ is the concentration of infected CD4$^+$ T cells that are actively producing mutant i, $v_i \in \mathbb{R}_+$ is the concentration of mutant virus i. R_i and N_i represent fraction bound ligand associated with infection and neutralization of mutants, respectively as described in Section 2.2.3. $\eta_c = 10^4$-10^5 is the number of CD4$^+$ T cell receptors per uninfected cell [51, 81]. We note that the infection binding function

$$R_i = \frac{[\text{gp}]^m}{[\text{gp}]^m + (K_{inf}^{wt}\Delta K_{inf}^{i})^m} \tag{2.8}$$

depends on the concentration of HIV glycoprotein for mutant i, and that is dependent on virus state v_i. The glycoprotein concentration for mutants i at time k is calculated by

$$[\text{gp}_i] = v_i[k]\frac{\eta_{gp120}}{N_A}, \tag{2.9}$$

where $\eta_{gp120} = 14$ is the number of glycoprotein molecules per virion [17, 54] and N_A is the Avogadro constant. d_x, d_y, d_v are degradation rates, respectively, of uninfected $CD4^+$ cells, infected $CD4^+$ cells, and virus. d_y and d_v are assumed to be equivalent for any mutant virus. λ is the T cell generation rate, and k_v is the viral burst rate, assumed to be equal for all mutants. Υ_i is the number of mutants i generated by the mutation process. Parameter values and units are listed in Table 2.4, Section 2.5.

Mutation process. The mutation process models the effects of error prone reverse transcription allowing the genetic variability necessary for selection. Experimental results indicate that single residue point mutations can cause resistance in the presence of even the most potent broadly neutralizing antibodies [22, 47, 101], Chapter 4. Thus, we focus our analysis on point mutations on the viral envelope glycoprotein but assume that mutations anywhere on the HIV genome are equally probable. We do not consider mutations based on insertions, deletions or recombinations.

We allow for any number of single nucleotide changes to occur based on the reverse transcription rate of mutation $u = 3 \times 10^{-5}$ mutations/base pair/replication cycle. However, we track viruses that have at most one residue substitution from wild type and assume that viruses with more than one residue substitution on contact sites are considered unfit and disappear. If a second residue change occurs outside the contact site, we assume this has the same fitness as the virus with one contact site residue change. Based on these criteria, any particular virus in the system can be in one of three states: wild type, point mutant with known fitness value and mutants with two

or more mutations. Mutation dynamics are represented by a Markov process.

Let $S = \{wt, m_1, ..., m_n, nf, uf\}$ be the state space associated with the Markov process, corresponding to infected cell types where wt is the wild type infected cell, $m_1, ..., m_n$ are point mutant infected cells that produce virus with known fitness, nf is an infected cell type that produces virus with unknown fitness, and uf is an infected cell type that produces virus considered unfit. As an approximation, we consider infected cells that produce mutant virus where no fitness knowledge to be equivalent in fitness to wild type, and those that produce virus with more than one residue change are considered unfit. Let X_n be a random variable denoting the state of a given cell at time n taking values in S, then its dynamics are given by the following state transition probabilities $P = \{p_{ij}\} = \mathbb{P}\{X_{n+1} = j | X_n = i\}$:

$$\begin{aligned}
p_{wt\to wt} &= (1-u)^{k_l} \\
p_{wt\to m_i} &= k_c u(1-u)^{k_c-1}(1-u)^{k_h} \\
p_{wt\to nf} &= k_h u(1-u)^{k_h-1}(1-u)^{k_c} \\
p_{wt\to uf} &= 1 - p_{wt\to wt} - p_{wt\to m_i} - p_{wt\to nf} \\
p_{m_i\to m_i} &= 1 - \textstyle\sum_{j\neq i}^{n} p_{m_i\to m_j} - p_{m_i\to uf} \\
p_{m_i\to m_j} &= \tfrac{2}{3k_c} p_{wt\to m_i} \\
p_{m_i\to uf} &= p_{wt\to nf},
\end{aligned} \tag{2.10}$$

for $i \in \{1, ..., n\}$ and $j \neq i$, and where k_l is the length of the entire HIV genome, k_c is the length of the genome for which fitness information exists, n_c is the number of possible point mutants with known fitness information, $k_h = k_l - k_c$, and u is the rate of mutation of HIV reverse transcriptase. All other state transition probabilities are not considered. This Markov process at the single cell level induces a Markov process at the population level.

2.4.2 Model Implementation and Simulations

Measuring the shape of viral distributions. The Gini coefficient measures statistical dispersion and is most commonly used as a measure of income inequality

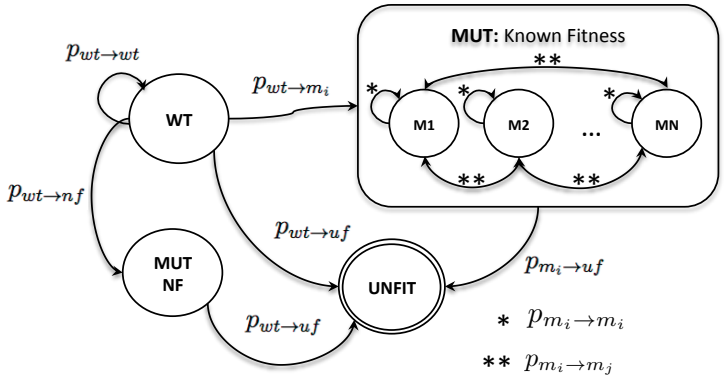

Figure 2.7: A finite state space Markov model representation of simplified mutation dynamics. MUTF is an abstract state that encompasses all point mutations for which fitness information is known, MUTNF is an abstract state that defines all mutations for which fitness information is not known. The WT state is a wild type virus state, UNFIT final state is a state that has no infectivity and is considered unfit with respect to any selective pressure.

in a population. We use this measure as an approximation to the shape of the virus distribution at steady state. The Gini coefficient G is calculated as

$$G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|y_i - y_j|}{2n^2\mu}, \tag{2.11}$$

where n is the number of all possible point mutations, y_i is the percent population of a point mutation i that occurs at steady state for all simulations and μ is their mean.

2.4.3 Experimental Methods

Mutagenesis, Virus Production and Cells. Site directed mutagenesis and assembly PCR were used to generate YU2-NL43 Env mutants. YU2-NL43 was modified using unique restriction sites EcoRI and XhoI. Inserts were generated by PCR using primers EcoRI-CF (5'-GCCAGCCAGAATTCTGCAACAACTGCTGTTTATCCATTTCAG-3') and XhoI-CR-(5'-GCGTCGACCTCGAGATACTGCTCCCACCCCATC-3') and individual sense and antisense mutagenesis primers corresponding to YU2-NL43 mutants listed in Figure 2.4. *Escherichia coli* One Shot STBL3 Chemically Competent cells (Life Technologies) were used to propagate proviral plasmids during

a 16 hour incubation at 30° C. Stocks were prepared using a DNA Midi kit (Zymo Research). All gene constructs were verified by complete sequencing of gp160. Cell-free virus was produced by transfection of HEK293T cells with YU2-NL43 virus coding plasmid using BioT (Bioland Scientific). Viral supernatant was harvested at 48 h post transfection, filtered through a 0.5 μm filter and aliquots were stored at -80° C. Stock concentrations were quantified by p24 enzyme-linked immunosorbent assay (ELISA) (Cell Biolabs). The YU2 Env/NL43 plasmid was obtained from the Nussenzweig lab, Rockefeller University. The green fluorescent protein (GFP) reporter T-cell line GXR-CEM is previously described in [11] was obtained through AIDS Research and Reference Reagent Program, National Institute of Allergy and Infectious Diseases, National Institutes of Health.

Protein Expression and Purification. Antibodies were transiently expressed in HEK293T/17 cells or suspension HEK 293-6E cells (National Research Council Biotechnology Research Institute, Montréal, QC, Canada) using 25-kDa linear polyethylenimine (Polysciences) for transfection. Supernatants were passed over MabSelect SuRe protein A resin (GE Healthcare) or Protein G Sepharose 4 Fast Flow (GE Healthcare) and eluted by using pH 3.0 citrate or glycine buffer, and then immediately neutralized. Antibodies were further purified by size exclusion chromatography using a Superdex 200 or 75 10/300 GL column.

In Vitro Replication and Neutralization Assays. To initiate infection for both replication and neutralization assays, GXR-CEM cells at 4×10^5 cells/ml were infected with 200 ng YU2-NL43 (HIV) virus stock in the absence of antibody and incubated for three days at 37° C. Two days after infection, uninfected GXR-CEM cells at 2×10^5 were pre-treated with 0, 0.02,0.08, 0.4,1.6, 6.4, 20, 80 μg/ml of NIH4546G54W antibody. Three days after initial infection with cell free virus, the infected GXR-CEM were washed and added to a final concentration of 1 % GFP-expressing donor cells to uninfected pre-treated GXR-CEM cells. For the neutralization assay, a constant concentration of antibody was maintained for each sample for three days following secondary infection. Infection was determined by measuring GFP reporter gene expression in the absence of antibodies measured daily for four days following sec-

ondary infection. Neutralization was determined by measuring the reduction in GFP reporter gene expression in the presence of antibodies NIH4546G54W for three days. Flow cytometry data was collected with a MACSQuant flow cytometer and transformed to a python format using the FlowCytometryTools python package (Gore lab, MIT). An automated gating and analysis software module written in python was developed to process large sets of flow cytometry data.

2.5 Appendix

Antibody	CVE	λ_1	λ_2	λ_3
3BNC117	0.387	5.05	3.25	1.7
3BNC55	0.3	5.05	4.6	0.35
NIH45-46G54W	0.612	1.45	1.45	7.1
45-46m2	0.46	1	4.15	4.85
45-46m7	0.425	2.35	6.4	1.25
VRC01	0.227	2.8	4.15	3.05
VRC03	0.316	30	60	10
VRC02	0.322	2.35	3.7	3.95
VRC-PG04	0.382	3.0	6.0	1.05
VRC-PG04b	0.382	50	40	10
b12	0.127	3.7	5.05	1.25
12A12	0.321	20	80	0
VRC-CH31	0.382	30	50	20
PG16	0.485	2.8	5.95	1.25
10-1074	0.373	1.9	3.25	4.85
10-996	0.486	4	5	1
PGT121	0.478	3.25	3.25	3.5
8ANC195	0.274	3.25	4.6	2.15
2F5	0.132	2.8	5.95	1.25
10E8	0.33	5.05	2.8	2.15
4E10	0.159	1.925	3.35	4.725

Table 2.1: 5-fold cross validation error (CVE), optimal λ_1, λ_2 and λ_3 for satlasso models on antibody neutralization data.

Antibody	Bound Virus	PDB accession
NIH-4546G54W	93TH057	4JKP
NIH-4546	93TH057	3U7Y
VRC-01	93TH057	3NGB
VRC-CH31	93TH057	4LSP
3BNC117	93TH057	4JPV
VRC-PG04	93TH057	3SE9
VRC-03	93TH057	3SE8
PG16	YU2	4DQ0
PGT128	HXB2	3TYG

Table 2.2: Antibody and virus molecular structures and their Protein Data Bank (PDB) accession numbers.

Ligand	Envelope molecule	K_d	Units	Ref
CD4	Core gp120	220	nM	[67]
CD4	full length gp120	22	nM	[67]
12A12	YU2-gp140	0.3	nM	[92]
3BNC60	YU2-gp140	6.81	nM	[92]
3BNC117	YU2-gp140	6.54	nM	[92]
3BNC55	YU2-gp140	57.8	nM	[92]
8ANC195	YU2-gp140	34.3	nM	[92]
45-46	YU2-gp140	6.75	nM	[92]
VRC01	YU2-gp140	0.4	mM	[92]
VRC-CH31	YU2 gp140	37.8	nM	[107]
VRC-PG04	KER2018 core	11.9	nM	[45]
VRC-03	YU2 gp140	16.1	nM	[104]
PG16	V1-V2 complex glycan	1.26	mM	[75]

Table 2.3: Ligand dissociation constants used in Equation 2.5.

Parameter	Value	Units
λ	10^5	cells $\cdot$ day^{-1}
d_x	0.01	day^{-1}
d_y	0.01	day^{-1}
k_v	100	day^{-1}
η_{CD4}	10^5	CD4 molecules per T cell
η_{gp120}	14	gp120 molecules per virion
τ		day
T	313	K
R	1.987	kcal $\cdot$ mol$^{-1}\cdot$ K^{-1}

Table 2.4: HIV evolutionary dynamics parameters and units corresponding to Equation (2.7) and Hill Equations (2.5).

Chapter 3

Robust Control of Evolutionary Dynamics

3.1 Introduction

A challenge inherent to the treatment of certain infectious and non-infectious diseases, such as HIV or cancer, is the risk that the pathogen or tumor will evolve away and become resistant to treatment methods that comprise the standard of care [35, 46, 48, 76]. Especially vulnerable to this phenomenon are treatment methods that involve exposing the disease population (such as viruses or cancer cells) to therapies targeting specific molecules involved in disease progression for an extended period of time. While these targeted therapies have the benefit of allowing physicians to tailor treatments to a patient's tumor cell population, they nonetheless establish an environment in which the occurrence of mildly drug resistant pathogens or tumor cells can develop an evolutionary advantage over those for which the therapy is targeted [23, 27, 41, 97], leading to so called "treatment-escape".

The challenge of designing treatment protocols that prevent escape is one that has been addressed by control theoretic methods. For cancer therapy, results in this spirit apply methods from optimal and receding horizon control [3, 16], as well as gain scheduling [2], to synthesize treatment protocols that are robust to parameter uncertainty, an inherent issue in all biological systems. Zhao et al. [105] present a static multi-objective optimization formulation to solve the combination therapy

problem for different initial tumor populations, when the drugs under consideration have additive, linear effects on cell viability. Proposed combination treatments were confirmed experimentally for different tumor initial conditions in a murine lymphoma model [106]. In the context of HIV and antiretroviral therapy, the authors in [25] propose a discrete time formulation that allows for the design of switching therapy strategies to delay the emergence of highly resistant mutant viruses. Recent results in [84] and [8], consider a simplified bilinear model and the optimal control problem is shown to be convex over a finite horizon for a predefined set of initial states.

This chapter presents three algorithms for the principled design of targeted combination drug treatment strategies that explicitly account for the evolutionary dynamics of a generic disease model, where the drugs under consideration are non-interacting and exhibit independent additive effects. Our first algorithm, introduced in [42], proposes a general iterative method that uses an $\mathcal{H}_\infty$ robust control approach to design targeted combination therapy concentrations and is effective in generating robustly stabilizing controllers. Our second algorithm addresses the lack of scalability symptomatic of semidefinite programming (SDP) formulations and proposes a scalable solution to the combination therapy problem by reformulating it as a second order cone program (SOCP), with robustness guarantees addressed by minimization of the induced $\mathcal{L}_1$ norm. We also require that the synthesized controller be not only robust to unmodeled dynamics, but also exhibit sparse structure and small feedback gains. This is motivated by the fact that the number of therapies commonly used in combination to treat a disease is often small while the number of potential usable therapies is often very large [1]. Targeted therapies such as small molecule drugs or antibodies exhibit a maximum effective concentration beyond which side effects are likely to worsen and no additional drug benefits are seen. Our third algorithm solves the combination therapy problem subject to the same design constraints (sparsity of the drug combination, maximum dosage and robustness constraints) formulated as an SOCP while addressing the nonlinear dynamics of individual drugs and of their combinations. In particular, through the piecewise linearization of individual and combination drug pharmacodynamics, in conjunction with a branch and bound like

algorithm for the effective search through these linear pharmacodynamic modes, we reduce the combination therapy problem to applying the SOCP formulation from [44] to a set of pharmacodynamic modes.

The chapter is organized as follows: in Section 3.2, we formulate the combination therapy control problem by introducing the evolutionary dynamics of a generic disease model with replication and mutation, subject to selective pressure from drugs. We also give a brief introduction to control of positive systems. Sections 3.3, 3.4 and 3.5 propose respectively, 1) an $\mathcal{H}_\infty$ robust control approach to design targeted combination drug treatment strategies, 2) an $\mathcal{L}_1$ scalable solution to the combination therapy problem, and 3) a $\mathcal{L}_1$ scalable solution to the combination therapy problem for a nonlinear pharmacodynamics model. We illustrate the applicability of these methods to antibody treatment strategies for HIV in Chapter 4.

3.2 Problem Formulation

We begin by fixing notation and introducing a simplified, general evolutionary dynamics model that encodes replication, mutation and drug selection. We observe that there is an inherent feedback structure to the aforementioned dynamics that forces us to consider *structured* controller synthesis. By showing that our dynamics are internally positive, we are able to leverage recent results in control of positive systems, that allow us to greatly simplify stability analysis and controller optimization with structure constraints.

3.2.1 Notation

$\mathbb{R}_+$ denotes the set of nonnegative real numbers. The inequality $X > 0, (X \geq 0)$ means that all elements of the matrix or vector X are positive (nonnegative). $X \succ 0$ means that X is a symmetric and positive definite matrix. The matrix $A \in \mathbb{R}^{n\times n}$ is said to be Hurwitz if all eigenvalues have negative real part. Finally, the matrix is said to be Metzler if all off-diagonal elements are nonnegative. Define $\mathbf{1}_n$ to be the vector of all ones of dimension n. The induced matrix norm for a matrix $M \in \mathbb{R}^{r\times m}$

is $\|M\|_{p-\text{ind}} = \sup_{w\in\mathbb{R}^m} \frac{|Mw|_p}{|w|_p}$ where $|w|_p = (|w_1|^p + ... + |w_m|^p)^{1/p}$. Let $G(s) = C(sI - A)^{-1}B + D$ be a $r \times m$ matrix transfer function. The induced norms of the corresponding impulse response $g(t) = Ce^{At}B + D\delta(t)$ are $\|g\|_{p-\text{ind}} = \sup_w \frac{\|g*w\|_p}{\|w\|_p}$ for $w \in \mathcal{L}_m^p[0,\infty)$, given that $g * w \in \mathcal{L}_p^r[0,\infty)$ is the convolution of g and w. Finally we refer to the ∞-induced robust controller as the $\mathcal{L}_1$ controller as is customary in the robust control literature [19].

3.2.2 Evolutionary dynamics model

The quasispecies model [26] was originally developed to describe the dynamics of populations of self replicating macromolecules undergoing mutation and selection. We choose this model for its relative simplicity and its ability to capture the salient features of the evolutionary dynamics of a simplified generic disease model. In this thesis, we incorporate the effects of potential therapies into the basic quasispecies model, by defining a drug binding reaction, $\ell + \rho \xrightarrow{K_A} \ell \cdot \rho$ - drug ℓ binds to self replicating macromolecule ρ with association rate K_A, giving a neutralized complex $\ell \cdot \rho$. The extended quasispecies model for n mutants and m drugs, is written as

$$\dot{x}_i = (r_i q_{ii} - d_i)x_i + \sum_{k\neq i}^{n} r_i q_{ik} x_k - \psi_i(\ell)x_i, \tag{3.1}$$

where $x_i \in \mathbb{R}_+$ is the concentration of mutant i, $\ell = (\ell_k) \in \mathbb{R}_+^m$ is the drug concentration (assumed to remain at constant concentrations throughout), r_i and d_i are the replication and degradation rates, respectively, of mutant i, and q_{ik} is the probability that mutant k mutates to mutant i (note that q_{ii} is the probability of no mutation occurring). The rates r_i can be viewed as the replication fitnesses of mutant i without the effect of the drug. When $\ell_k = 0$, for all $k \in \{1, ..., m\}$, the quasispecies dynamics are unstable. Finally, the function $\psi_i(\ell)$ represents the pharmacodynamics of individual drugs ℓ_k and their combinations with respect to the i-th mutant species. The map $\psi_i(\ell)$ is a function defined as the sum of nonlinear functions representing additive drug effects, represented by Hill equations and described in the next section.

3.2.3 The Hill equation

The Hill equation has been used in pharmacology to model nonlinear drug dose-responses, such as the concentration dependent effects of drugs on cell viability or virus neutralization. More generally it serves to quantify drug-receptor relationships, the fraction of bound receptors ρ (e.g. cell receptors, virus proteins) as a function of ligand ℓ_k (e.g. drug, antibody) concentrations for the binding reaction

$$\begin{aligned} &\ell_k + \rho \xrightarrow{K_A} \ell_k \cdot \rho, \\ &\psi_\rho(\ell_k) = \tfrac{[\ell_k]^{n_k}}{[\ell_k]^{n_k} + K_k^{n_k}}, \end{aligned} \tag{3.2}$$

where $\psi_\rho(\ell_k) \in [0, 1]$ is the fraction of bound receptors, $\ell_k \in \mathbb{R}_+$ is the concentration of ligands, $K_k = \frac{1}{K_A} \in \mathbb{R}_+$ is the dissociation constant associated with the binding reaction (3.14), $n_k \in \mathbb{R}_+$ is the Hill coefficient that represents the degree of cooperativity, i.e. the degree to which binding of a ligand molecule modulates the probability of another ligand molecule binding.

We notice that the Hill function is a biological analog to actuator saturation, in that there is a law of diminishing returns in terms of the effect of ever increasing drug concentrations on the system. In particular, the Hill functions $\psi_\rho(\ell_k)$ that appear in most pharmacodynamic models look approximately linear for small to moderate ℓ_k, and nearly constant for large ℓ_k. The application of algorithms developed in previous Sections 3.3 and 3.4 assume that the Hill equation can be approximated by a linear function and inequality constraints. In Section 3.5, we provide a finer approximation of the Hill function by using piecewise linear approximations to solve the combination therapy problem subject to nonlinear pharmacodynamics. For all algorithms developed in this chapter, we consider combinations of independently acting drugs that exhibit an additive effect.

3.2.4 State space representation

The following state space representation of equation (4.6) emphasizes the inherent feedback structure that arises from drug binding reactions:

$$\begin{aligned} \dot{x} &= (A - \Psi(\ell))x + w \\ z &= Cx \end{aligned} \tag{3.3}$$

with $A \in \mathbb{R}^{n\times n}$, with $A_{ij} = r_i q_{ij} \geq 0 \; \forall \; i \neq j$ and $A_{ii} = r_i q_{ii} - d_i$, that encodes the mutation and replication dynamics; $\Psi(\ell) \in \mathbb{R}^{n\times n}$, a diagonal matrix, with diagonal elements $\psi_i(\ell)$ representing additive drug effects for each mutant i; $\ell = (\ell_k) \in \mathbb{R}^m$, a vector of individual drug concentrations; $C = \mathbf{1}_n^T \in \mathbb{R}^{1\times n}$; and $w \in \mathbb{R}_+^n$ an arbitrary positive disturbance. We set the regulated output $z = \mathbf{1}_n^T x$ to be the total mutant population, so as to ensure that the resulting treatment plan is one that robustly drives the total mutant population to zero.

If we approximate drug effects for each mutant $\psi_i(\ell)$ be a linear function of ℓ, then we can construct an appropriately defined block diagonal matrix Ψ such that $\Psi(\ell) = \Psi L$ where $L = I \otimes \ell \in \mathbb{R}^{mn\times n}$. In this particular case, the closed loop transfer function from input to output system is given by

$$G(s) = C(sI - (A - \Psi L)^{-1}B + D. \tag{3.4}$$

The control task then becomes to engineer drug concentrations ℓ by finding a "controller" $L = I \otimes \ell \in \mathbb{R}^{mn\times n}$ that leads to a stable G satisfying $||G||_{\infty-\text{ind}} < \gamma$, for some desired robustness level $\gamma > 0$.

3.2.5 Control of positive systems

Recent results on the synthesis of controllers for positive systems show that the design of structured static state feedback controllers for internally positive systems can be reformulated as a convex problem with methods that scale *linearly* with the number of non zero elements in the feedback matrix [82, 83].

Briefly, we recall definitions from positive systems theory [96]. We will then expand the background on control of positive systems as needed in Section 3.3.1 and Section 3.4.1.

Definition 2. *The LTI system*

$$\begin{aligned} \dot{x} &= Ax + Bw \\ z &= Cx + Dw \end{aligned} \tag{3.5}$$

with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times q}$, $C \in \mathbb{R}^{p\times n}$ *and* $D \in \mathbb{R}^{p\times q}$ *is called* internally positive *if for every* $x_0 \in \mathbb{R}^n_+$ *and all inputs such that* $w(t) \in \mathbb{R}^q_+$ *for all* $t \geq 0$*, the state vector* $x(t)$ *belongs to* $\mathbb{R}^n_+$ *and the output vector* $z(t)$ *belongs to* $\mathbb{R}^p_+$ *for all* $t \geq 0$*.*

The internal positivity of a system is easily determined by a simple condition on its system matrices:

Lemma 3. System (3.5) is internally positive if and only if

1. A is Metzler, and
2. $B, C, D \geq 0$, i.e. matrices B, C, and D are entry-wise non-negative.

We observe that in the state space system (3.3), A is a Metzler matrix with off-diagonal entries that are several orders of magnitude smaller than the diagonal entries. This is due to the biological fact that mutation rates range from 10^{-5} - 10^{-8} mutations per base pair per replication cycle for reverse transcriptase [58] to DNA replication [68].

In light of this result, it is straightforward to show that system (3.3) is internally positive:

Lemma 4. System (3.3) is internally positive.

Proof. Condition 2) of Lemma 3 is easily seen to be satisfied by noting that in equation (3.3), $B = I$, $C = \mathbf{1}^T$ and $D = 0$. To see that $A - \Psi(\ell)$ is Metzler, it suffices to notice that since $\Psi(\ell)$ is strictly diagonal, it cannot affect the Metzler property of A. □

3.3 Static state feedback strategies for combination therapy using H_∞ control

We introduce a general algorithm for the systematic design of feedback strategies to stabilize the evolutionary dynamics of a generic disease model using an $\mathcal{H}_\infty$ approach. We show that designing therapy concentrations can be cast as an $\mathcal{H}_\infty$ state feedback synthesis problem, where the feedback gain is constrained to not only be strictly diagonal, but also that its diagonal elements satisfy an overdetermined set of linear equations. Leveraging recent results in positive systems [82, 96], we develop an algorithm that *always* yields a stabilizing controller.

We recall recent results on the synthesis of structured static state feedback controllers for positive systems in Section 3.3.1, and propose an algorithm to solve for a stabilizing H_∞ sub optimal controller in Section 3.3.3.

3.3.1 The bounded real lemma for internally positive systems

Recent results by Tanaka and Langbort [96] and Rantzer [82] on the synthesis of $\mathcal{H}_\infty$ controllers for positive systems show that the design of structured static state feedback controllers for internally positive systems can be reformulated as a convex problem. In this section we provide a brief survey of the relevant definitions and results from [96].

In Section 3.2.5, we showed that our system (3.3) is internally positive. Systems that are internally positive enjoy the significant advantage that the storage function matrix used in the bounded real lemma to characterize the $\mathcal{H}_\infty$ norm of a system via a semi-definite program (SDP), can be taken without loss to be diagonal, as outlined in the following theorem, slightly modifed from [96].

Theorem 5. *Let the system (3.5) be internally positive with (A, B) stabilizable and (C, A) detectable. Let the corresponding transfer function be given by $G(s) := C(sI - A)^{-1}B + D$. Then the following statements are equivalent:*

1. *$\|G\|_\infty < \gamma$ and A is Hurwitz;*

2. There exists a diagonal matrix $X > 0$ such that

$$\begin{bmatrix} A^TX + XA + C^TC & XB + C^TD \\ B^TX + D^TC & D^TD - \gamma^2 I \end{bmatrix} \prec 0. \tag{3.6}$$

The fact that X can be restricted to be diagonal is very useful in synthesizing *structured* feedback controllers, when this structure is defined by *sparsity* in the feedback gain. Our setting, however, requires not only sparsity, but a type of algebraic consistency — the controller structure $L = I \otimes \ell$ implies that each block diagonal element of L must be equal. Unfortunately, there is no known way of enforcing this additional coupling in a convex manner.

In the following sections, we deal with the aforementioned non-convexity of the optimal control problem by formulating an iterative algorithm for finding effective drug concentrations, exploiting the internal positivity of the system to show that it always yields a stabilizing controller.

3.3.2 Initializing stabilizing controller

We begin by synthesizing a stabilizing controller to use as an initial controller in our iterative algorithm, as noted, which admits a particularly simple formulation in light of the Metzler nature of A and the diagonal structure of ΨL.

Lemma 6. There exists $\epsilon > 0$ such that the solution to the convex program

$$\begin{aligned} &\text{minimize } _{\ell \in \mathbb{R}^m_+} ||\ell||_\infty \\ &\text{subject to} \\ &A_d + \epsilon I - \Psi L \prec 0 \\ &L = I \otimes \ell \end{aligned} \tag{3.7}$$

is a stabilizing controller for system (3.3), where A_d is a diagonal matrix comprised of the diagonal elements of A.

Proof. Rewrite $A = A_d + M$ where A_d is diagonal and $M = \{m_{ij}\} \in \mathbb{R}^{n \times n}$, $m_{ij} = 0$ for $i = j$ and $m_{ij} > 0$ for $i \neq j$. By the Perron-Frobenius theorem, there exists $r > 0$ such that the spectral radius $\rho(M) = r \leq \max_i \sum m_{ij}$. Let $\epsilon = \max_i \sum m_{ij}$

and rewrite $M = \epsilon I - (\epsilon I - M)$. We note that $-(\epsilon I - M) \prec 0$. The closed loop dynamics are then given by $A - \Psi L = A_d + \epsilon I - (\epsilon I - M) - \Psi L \prec A_d + \epsilon I - \Psi L \prec 0$, yielding the desired stability. □

Remark 7. The stabilization problem can be solved independently of a storage function because it can be reduced to satisfying element wise inequalities.

3.3.3 An $\mathcal{H}_\infty$ combination therapy controller

Observe that through a straightforward application of inequality (3.6) to system (3.3), the antibody concentrations ℓ yielding an optimal $\mathcal{H}_\infty$ closed loop norm can be found by solving the following non-convex program:

$$\begin{aligned}
&\text{minimize } \gamma \\
&\text{subject to} \\
&\begin{bmatrix} A_{cl}^T X + X A_{cl} + (\overline{\Psi L})^T (\overline{\Psi L}) & X \\ X & -\gamma^2 I \end{bmatrix} \prec 0 \\
&A_{cl} = (A - \Psi L) \\
&L = I \otimes \ell \\
&X \succ 0,\ X \text{ diagonal}
\end{aligned} \tag{3.8}$$

Applying a Schur complement to $A_{cl}^T X + X A_{cl} + (\overline{\Psi L})^T (\overline{\Psi L})$ yields the more amenable form

$$\begin{aligned}
&\text{minimize } \gamma \\
&\text{subject to} \\
&\begin{bmatrix} A_{cl}^T X + X A_{cl} & X & (\overline{\Psi L})^T \\ X & -\gamma I & 0 \\ \overline{\Psi L} & 0 & -\gamma I \end{bmatrix} \prec 0 \\
&A_{cl} = (A - \Psi L) \\
&X \succ 0,\ X \text{ diagonal}
\end{aligned} \tag{3.9}$$

Remark 8. We can impose an additional constraint limiting the concentrations of candidate therapies. This is necessary with certain drugs that have maximum tolerated doses dictated by clinical trials.

Thus the only non-convexity remaining are the product terms between the storage function matrix X and the controller gain ℓ in $A_{cl}^T X + X A_{cl}$. As mentioned earlier, there are no known convex reformulations of this problem due to the additional structure on ℓ. As such, we suggest the following iterative algorithm, based on the convex programs (3.7) and (3.9), to find a stabilizing controller.

For ease of notation, let $\mathrm{P}_{X'}(\ell, \gamma)$ denote that we solve (3.9) with $X = X'$ fixed, and that we optimize over ℓ and γ. Similarly, let $\mathrm{P}_{\ell'}(X, \gamma)$ denote that we solve (3.9) with $\ell = \ell'$ fixed, and that we optimize over X and γ. Additionally, let $(Z, \gamma) = \mathrm{P}_{Z'}(Z, \gamma)$ denote the solutions to the respective programs, for $Z, Z' \in \{X, \ell\}$. We are now in a position to present our algorithm:

Algorithm 1 $\mathcal{H}_\infty$ Combination Therapy

1. Set $\epsilon > 0$
2. Solve (3.7) to obtain an initial stabilizing controller ℓ'.
3. while $\gamma' - \gamma > \epsilon$:
   (a) Set $(X', \gamma) = \mathrm{P}_{\ell'}(X, \gamma)$.
   (b) Set $(\ell', \gamma) = \mathrm{P}_{X'}(\ell, \gamma)$.
   (c) Set $\gamma' = \gamma$.

Proposition 9. *Algorithm 1 always converges to a feasible γ and generates a stabilizing controller for (3.3).*

Proof. By Lemma 6, an initial stabilizing controller can always be found, and thus the algorithm can always be initialized. The sequence of γs then defined by the iterative process in Algorithm 1 is non-increasing by construction, and bounded below by 0, thus implying convergence. We therefore have that our algorithm always converges

to a local minimum value of γ, and yields a set of gains which, by the bounded real lemma, robustly stabilizes system (3.3). □

3.4 Static state feedback strategies for combination therapy for large-scale systems

In Section 3.3, we introduced a general algorithm that used an $\mathcal{H}_\infty$ approach for the principled design of targeted combination therapy concentrations that explicitly account for the evolutionary dynamics of a generic disease model. This algorithm was effective in generating robustly stabilizing controllers, however it suffered from an inherent lack of scalability symptomatic of semidefinite programming formulations. Here,we propose a scalable solution to the combination therapy problem by reformulating it as a second order cone program (SOCP). We simultaneously address the requirement that the synthesized controller be not only robust to unmodeled dynamics but also exhibit sparse structure and small feedback gains, and allow the designer to explore respective tradeoffs. In particular, through ℓ_1 and ℓ_2 regularization, we induce sparse structure in the feedback controller while bounding the magnitude of the feedback gains, leading to a SOCP formulation of the combination therapy synthesis problem. In this section, we propose a scalable algorithm for the systematic design of sparse, small gain feedback strategies to stabilize the evolutionary dynamics of a generic disease model.

3.4.1 Controller synthesis by linear programming

In this section, we provide a brief survey of the relevant definitions and results from design of structured static state feedback controllers for positive systems, [83].

Theorem 10. *For the system*

$$\begin{aligned} \dot{x} &= (A + ELF)x + Bw \\ z &= (C + GLF)x + Dw \end{aligned} \tag{3.10}$$

let $\mathcal{D}$ be the set of $m \times m$ diagonal matrices with entries in [0,1]. Suppose that $A+ELF$ is Metzler and $C + GLF \geq 0$ for all $L \in \mathcal{D}$. Let $g(t)$ be the impulse response of

$$G(s) = (C + GLF)[sI - (A + ELF)]^{-1}B + D$$

If the matrices B, D and F have non-negative coefficients, then the following two conditions are equivalent:

1. *There exists an $L \in \mathcal{D}$ with $A + ELF$ Hurwitz and $\|g\|_{\infty-\text{ind}} < \gamma$.*
2. *There exists a $\xi \in \mathbb{R}^n_+, \mu \in \mathbb{R}^m_+$ with*

$$A\xi + E\mu + B\mathbf{1} < 0$$

$$C\xi + G\mu + D\mathbf{1} < \gamma\mathbf{1}$$

$$\mu \leq F\xi$$

If ξ, μ satisfy the linear constraints 2) then the stability and norm guarantees of 1) hold for every L such that $\mu = LF\xi$.

Input-output performance is characterized using induced norms, which are determined by the closed loop system's static gain:

Theorem 11. *For an $r \times m$ transfer matrix $G(s) = C(sI - A)^{-1}B + D$, let $g(t) = Ce^{At}B + D\delta(t)$ be the corresponding impulse response, where $Ce^{At}B \geq 0$ for $t \geq 0$ and $D \geq 0$, with A Hurwitz. Then $\|g\|_{p-\text{ind}} = \|G(0)\|_{p-\text{ind}}$ for $p = 1$, $p = 2$ and $p = \infty$.*

The positive nature of the system allows us to restrict ourselves to linear storage functions, which in turn allows for sparse structure to be imposed on the feedback gain [82]. Our feedback gain $L = I \otimes \ell$ is not only structurally constrained to be block diagonal, but algebraically constrained as well, in that all block diagonal components must be equal. Unfortunately, there is no known convex reformulation for this additional constraint.

3.4.2 Regularization for structured controller synthesis

The biomedical justification for wanting a simple controller structure with small gains is twofold: first, the number of therapies that can be used simultaneously to treat a disease is often limited, and second to minimize side effects, it is desirable to keep the magnitude of drug concentrations small while being robust to pharmacodynamic perturbations. These design specifications can be expressed with the use of regularization, a common technique used in machine learning and inverse problems for *model identification* [13, 24, 85, 15] and increasingly used for controller *design* [60, 53, 21, 28, 59]. As such, we introduce ℓ_1 and ℓ_2 penalties in our design objective to promote controller sparsity and minimize controller gains.

We combine these regularization techniques with controller synthesis results for positive systems and present an iterative algorithm that yields a suboptimal $\mathcal{L}_1$ controller. This formulation of the combination therapy problem allows the designer to explore explicit trade offs between closed loop performance, sparsity in controller structure and gain minimization.

In the following sections, we address the aforementioned non-convexity of the optimal control problem by formulating an iterative algorithm for finding effective drug concentrations. Our main result addresses the issue of synthesizing a stabilizing controller subject to the constraints imposed by the quasispecies model (3.3), with acceptable robustness properties characterized in terms of its ∞-induced closed loop norm.

3.4.3 A scalable $\mathcal{L}_1$ combination therapy controller

Observe that through a straightforward application of Theorem 11, with $B = I,\ C = \mathbf{1}^T,\ D = 0,\ E = -\Psi,\ F = I$ to system (4.6), solving the following non-convex program:

$$\begin{aligned}
&\underset{\ell\in\mathbb{R}^m_+, x\in\mathbb{R}^n_+}{\text{minimize}} \ \|Cx\|_\infty \ + \lambda_1\|\ell\|_1 + \lambda_2\|\ell\|_2 \\
&\text{subject to} \\
&\quad Ax + Kx + \mathbf{1} \leq 0 \\
&\quad K = \Psi L \\
&\quad L = I \otimes \ell \\
&\quad x \geq 0
\end{aligned} \tag{3.11}$$

will yield a sparse combination of drug concentrations ℓ, yielding an optimal ∞-induced closed loop norm for appropriately chosen regularizers $\lambda_1 \geq 0, \lambda_2 \geq 0 \in \mathbb{R}$.

Remark 12. We can impose an additional constraint limiting the concentrations of candidate therapies. This is necessary with certain drugs that have maximum tolerated doses dictated by clinical trials.

As mentioned earlier, there are no known convex reformulations of this problem due to the additional structure on L. As such, we suggest the following iterative algorithm, based on the convex programs (3.9) and (3.23), to find a stabilizing controller. For notation, let $Y = \mathrm{P}_Z(x, \gamma)$ denote an optimization problem P in which we optimize over x and γ leaving Z fixed and with solution Y.

Program 1. $\mathrm{P1}_\ell(x, \gamma)$:

$$\begin{aligned}
&\underset{\gamma, x\in\mathbb{R}^n_+}{\text{minimize}} \ \gamma \\
&\text{subject to} \\
&\quad Ax + \Psi Lx + \mathbf{1} \leq 0 \\
&\quad L = I \otimes \ell, \ \ C = \mathbf{1}^T \\
&\quad x \geq 0 \\
&\quad \gamma \geq \|Cx\|_\infty
\end{aligned} \tag{3.12}$$

Program 2. $\mathrm{P2}_{(x,\lambda_1,\lambda_2)}(\ell, \gamma)$

$$
\begin{aligned}
&\underset{\gamma,\ell\in\mathbb{R}^m_+}{\text{minimize}}\ \gamma + \lambda_1\|\ell\|_1 + \lambda_2\|\ell\|_2 \\
&\text{subject to} \\
&\quad Ax + \Psi Lx + \mathbf{1} < 0 \\
&\quad L = I \otimes \ell,\ \ C = \mathbf{1}^T \\
&\quad \gamma \geq \|Cx\|_\infty
\end{aligned}
\tag{3.13}
$$

We now present our algorithm:

Algorithm 2 Scalable Combination Therapy

1. Set $\epsilon > 0$.
2. *Solve for initial stabilizing controller* ℓ':

   Solve convex program (3.7) to obtain controller ℓ_0.
   Set $(x', \gamma) = \mathrm{P1}_{\ell_0}(x, \gamma)$.
   Set $(\ell', \gamma) = \mathrm{P2}_{(x',0,0)}(\ell, \gamma)$.
3. *Find* $(\lambda_1', \lambda_2', \ell)$ *that minimize* γ:

   $\forall(\lambda_1, \lambda_2) \in \Lambda_1 \times \Lambda_2, \Lambda_1, \Lambda_2 \in \mathbb{R}^k_+$,
   while $\gamma' - \gamma > \epsilon$:
   Set $(x', \gamma) = \mathrm{P1}_{\ell'}(x, \gamma)$.
   Set $(\ell', \gamma) = \mathrm{P2}_{(x',\lambda_1,\lambda_2)}(\ell, \gamma)$.
   Set $\gamma' = \gamma$.

Remark 13. The sequence of γ's defined by the iterative process in Algorithm 1 is non increasing

Remark 14. In practice, we note that the $\mathcal{L}_1$ controller suffers from dependence on initial conditions and converges to local optima quickly, yielding a stabilizing controller with robustness properties that are not significantly different from the nominal controller. A solution to this is to iterate once through P1 and P2, with $\lambda_1 = \lambda_2 = 0$ and initialize the algorithm with the resulting controller.

Remark 15. Due to the presence of the ℓ_2 regularizer, P2 and (5) are SOCPs and not linear programs. As it will be clearly demonstrated in an example in the next chapter, this is still more efficient than the SDP combination therapy algorithm from

Section 3.3.3 [42]. In addition, the second order cone constraint is only on the drug concentrations so has a minimal effect on performance [].

3.5 Feedback strategies for combination therapy for large scale systems with nonlinear pharmacodynamics

In Section 3.4, we introduced a scalable, iterative algorithm for the principled design of targeted combination therapy concentrations that explicitly accounts for the evolutionary dynamics of a generic disease model. This algorithm was effective in generating robustly stabilizing controllers, while simultaneously allowing the designer to explore explicit trade offs between closed loop performance, sparsity in controller structure and gain minimization. Leveraging recent results from positive systems [83], we formulated this algorithm as a second order cone program (SOCP), which made the controller synthesis scalable. However it could not take into account a) the pharmacodynamics of the input, potentially suffering from over or underestimating gains, and b) the effects of synergistic or antagonistic drug combinations that can be modeled with additional nonlinear pharmacodynamic terms [102].

In this section, we propose an algorithm that solves the combination therapy problem subject to the same design constraints (sparsity of the drug combination, maximum dosage and robustness constraints) formulated as an SOCP while addressing the nonlinear dynamics of individual drugs and of their combinations. In particular, through the piecewise linearization of individual and combination drug pharmacodynamics, in conjunction with a branch and bound like algorithm for the effective search through these linear pharmacodynamic modes, we reduce the combination therapy problem to applying the SOCP formulation from [44] to a set of pharmacodynamic modes.

The main contribution of this section is an algorithm for the systematic design of *sparse*, *small gain* feedback strategies to stabilize the evolutionary dynamics of a

generic disease model and general nonlinear pharmacodynamics models, which support synthesis of feedback strategies in light of highly nonlinear drug dynamics.

3.5.1 Pharmacodynamic models

Pharmacodynamic models are used to quantitatively describe nonlinear drug dose-responses and model drug-receptor relationships. These models are often described using combinations of nonlinear Hill functions, and allow for the modeling of drug saturation effects [86, 90, 78]. These pharmacodynamic nonlinearities are further increased with the fact that drugs administered in combination can have independent additive effects or can otherwise exhibit synergistic or antagonistic dose dependent behavior, further complicating the design of suitable combination therapies [102]. There have been several attempts to deal with nonlinear HIV infection dynamics using model predictive control (MPC) to design optimal antiretroviral drug dosing strategies [56, 108]. With MPC and other linear control synthesis methods, an essential feature of the system, that of nonlinearities associated with drug binding, is linearized away. In particular, such an approach can lead to a model that underestimates the efficacy of a drug at lower concentrations, and over estimates its efficacy at high dosages *unless* a sufficiently small update time is taken. This latter restriction may then lead to strategies that are no longer realistic in a biomedical application, where it may be difficult to update the administered therapies at the frequency dictated by the MPC controller. In this section, we take an alternative approach and attempt to design constant drug therapies by taking input nonlinearities into account more explicitly. In particular, through the piecewise linearization of individual and combination drug pharmacodynamics, in conjunction with a branch and bound like algorithm for the effective search through these linear pharmacodynamic modes, we reduce the combination therapy problem to applying the SOCP formulation from [44] to a set of pharmacodynamic modes.

When combined, drugs can have additive, synergistic or antagonistic effects that need to be explicitly accounted for when designing combination therapies. When the

presence of one drug modulates the effect of another, the combined drug effects are no longer additive, and in particular, the phamacodynamics of the drug interaction now incorporate additional terms that represent this interaction.

We use the Hill equation introduced in Section 3.2.3 to model the nonlinear pharmacodynamics of independently acting drugs, in which the effect of each drug on the system is additive. Consider the system of m drug binding reactions to different receptors on a particular cell or virus x:

$$\ell_1 + \rho_1 \xleftarrow{K_1} \ell_1 \cdot \rho_1, ..., \ell_m + \rho_m \xleftarrow{K_m} \ell_m \cdot \rho_m \tag{3.14}$$

where ℓ_k is the drug k and ρ_k is a receptor. If these receptors comprise different drug binding targets on a cell or virus x, then we can describe the total effect of these independently acting drugs $\ell_1, ..., \ell_m$ on x as

$$\Psi_x(\ell) = \sum_{k=1}^{m} \psi_{\rho_k}(\ell_k). \tag{3.15}$$

Remark 16. Although we focus on this drug interaction model, we note that our approach applies nearly verbatim to a synergistic/antagonistic drug binding model – it suffices to use a suitable expression for $\Psi_x(\ell)$ to take these interactions into account. We do note however that this may lead to a more involved piecewise linear approximation procedure.

In the following sections, we formulate the task of designing suitable combination therapies as an optimal control problem. The inherent nonlinearities of the system make this a challenging task — in [42, 44], we worked with a simplified problem in which we assumed the $\Psi_x(\ell)$ were linear functions – in this section we relax that assumption, and show that at the expense of some additional modeling complexity, we are able to reduce the problem to that considered in [44]. Our main result is based around the use of piecewise linear approximations, and relating the robustness levels of the approximate system to that of the true underlying system.

3.5.2 Piecewise linear mode approximations and mode reduction

In the following sections, we assume that each $(\Psi(\ell))_{ii} = \Psi_{x_i}(\ell)$ has the form given by Equation (3.14). To take into account nonlinear pharmacodynamics, we propose a piecewise linear approximation algorithm and a mode reduction algorithm for problems where there is a large number of non interacting or synergistic drug combinations.

We assume that the pharmacodynamics for every individual drug and combination are defined over the same drug concentration domain $\mathcal{D} \in \mathbb{R}$. Let $\mathcal{P} = \{p_1, ..., p_k\}$ be a partitioning of this domain into k intervals.

Definition 17. *A pharmacodynamic mode $\omega = (\omega_1, ..., \omega_m)$ is an m-tuple in $\mathcal{P}^m$. The total number of pharmacodynamic modes is $|\mathcal{P}|^m$ where m is the number of drugs under consideration. For $\ell \in \mathbb{R}^m$, we define $\ell \in \omega \iff \ell_i \in \omega_i$, for all $i = \{1, 2, ..., m\}$.*

We let $\psi_{i\omega} : \mathbb{R}^m \to \mathbb{R}$ be the affine approximation of the pharmacodynamics for each mutant i for $\ell \in \omega$, i.e. the sum of the individual and combination drug effects on mutant i while operating within mode ω. We can then construct a linear approximation to $\Psi(\ell)$ via an appropriately defined block diagonal matrix Ψ_ω, constructed from the $\psi_{i\omega}$ (c.f. [44]), and write $\Psi(\ell) = \Psi_\omega L_\omega$, where $L_\omega = (I \otimes \ell_\omega) \in \mathbb{R}_+^{mn \times n}$ is the block diagonal matrix, with identical block diagonal elements given by the drug concentrations $\ell \in \omega$. The resulting dynamics, for a fixed concentration $\ell \in \omega$, are then described by the transfer function

$$G_{\ell_\omega}(s) = C(sI - (A - \Psi_\omega L_\omega)^{-1} B + D. \tag{3.16}$$

We consider the problem of finding a suitably sparse therapy combination that achieves a certain closed loop performance level γ. As such, our initial goal becomes to reduce the search space to the set of sparse modes that achieve the desired level of robustness, where a sparse mode ω is one that allows at least one drug concentration

to be zero, in other words, $0 \in \omega_i$ for at least one $\omega_i \in \omega$.

In order to do so, we require two lemmas. The first provides sufficient conditions on the linear approximation terms ψ_{iw} that guarantee that the robustness of the piecewise linear approximation is an upper bound on that of the true system. The second is the simple observation that for non-interacting or synergistic additive drug interactions, the robustness of the closed loop dynamics increases as drug concentrations are uniformly increased (this statement will be made precise). This result allows us to subsequently develop a branch and bound like method that significantly reduces the search space of the algorithm.

We begin with a result on the input-output performance of a positive system, taken from [83].

Lemma 18. Let $G(s) = C(sI - A)^{-1}B + D$ be a positive system. Then $\|G\|_{\infty-\text{ind}} \leq \gamma$ if and only if there exits $x \geq 0$ such that

$$\begin{aligned} Ax + B\mathbf{1}_n &< 0 \\ Cx &\leq \gamma \end{aligned} \tag{3.17}$$

Lemma 19. Let $\Psi(\ell)$ be the nonlinear pharmacodynamics function, $\psi(\ell)$ the vector of its diagonal elements, and denote by $\Psi_\omega(\ell)$ its piecewise linear approximation within mode ω, and by $\psi_\omega(\ell)$ the vector of its diagonal elements. If for every mode ω we have $\psi_\omega(\ell) \leq \psi(\ell)$ for all $\ell \in \omega$ then the $\mathcal{L}_1$ norm γ of the piecewise linear approximation (3.16) is an upper bound on that of the true system (3.3).

Proof. Note that for a fixed ℓ, both the full and piecewise linear approximation systems are linear in x. Therefore, by Lemma 18, the $\mathcal{L}_1$ norm of (3.16) is upper bounded by γ if and only if there exists an $x > 0$ such that

$$\begin{aligned} (A - \Psi_\omega(\ell))x + \mathbf{1}_n &< 0 \\ \mathbf{1}_n^T x &\leq \gamma \end{aligned} \tag{3.18}$$

Thus it suffices to show that this same x yields $(A - \Psi(\ell))x + \mathbf{1}_n \leq 0$ and the desired

conclusion follows immediately. To that end, rewrite the first inequality of (3.18) as

$$(A - \Psi(\ell))x + \mathbf{1}_n < (\Psi_\omega(\ell) - \Psi(\ell))x, \tag{3.19}$$

and notice that the right hand side is less than or equal to 0 by the assumptions of the lemma. □

This lemma essentially states that if our piecewise linear approximation is conservative, then the norm of the approximate system serves as a certificate for the norm of the true system.

Next we formalize the observation that increasing the concentrations of non or synergistically interacting therapies present in the system will improve robustness.

Lemma 20. Let ℓ^1 and ℓ^2 be therapy combinations such that $\ell^1 \geq \ell^2$. Then if the piecewise linear approximation $\Psi_\omega(\ell)$ is non-decreasing, $||G_1||_{\infty-ind} \leq ||G_2||_{\infty-ind}$, where

$$G_i = C(sI - (A - \Psi_\omega(\ell^i))^{-1}I.$$

Proof. By assumption, the piecewise linear function $\Psi_\omega(\ell)$ is non-decreasing. Thus if $\ell^1 \geq \ell^2$, then $\Psi_\omega(\ell^1) \geq \Psi_\omega(\ell^2)$. Let $\gamma_i = ||G_i||_{\infty-ind}$. By Lemma 18, γ_i is the solution to the following optimization:

$$\begin{aligned} &\underset{\gamma, x \geq 0}{\text{minimize}}\ \gamma \\ &\text{subject to} \\ &(A - \Psi_\omega(\ell))x + \mathbf{1} < 0 \\ &Cx \leq \gamma \end{aligned} \tag{3.20}$$

Let γ_2 and x be the optimal solutions of the above program for $i = 2$. Then we have that

$$(A - \Psi_\omega(\ell^1))x + \mathbf{1} \leq (A - \Psi_\omega(\ell^2))x + \mathbf{1} \leq 0. \tag{3.21}$$

Hence (γ_2, x) is a feasible solution for optimization (3.20) with $i = 1$, implying that $\gamma_1 \leq \gamma_2$. □

Thus we see that by constraining the piecewise linear approximations to be under approximations of the effects of the drugs, we are able to bound the performance of the true system.

Remark 21. We note that these results typically will not hold for antagonistic drug interactions — however, many settings in which such models are required (such as cancer therapy design) do not have a large set of therapies or mutants, thus mitigating the computational cost of the sequential search across modes.

We exploit this result to reduce the modes that need to be searched over — in particular, we use the partial order implied by the previous two results to upper and lower bound the uniform concentration treatments required to achieve a prescribed performance level γ. We also exploit the fact that we are searching for *sparse* treatment strategies to further eliminate modes.

This approach is formalized in the following algorithm. Let $\ell_{\omega_{max}} \in \mathbb{R}^m$, be the maximum possible drug concentrations.

Algorithm 3 Sparse mode reduction algorithm

Set $\ell_{\omega_i} \leftarrow \ell_{\omega_{max}}, \gamma > 0$.

while $\neg(\ell_{\omega_i} == \mathbf{0}_n \ \wedge \ \epsilon)$:

if $||G_{\ell_{\omega_i}}|| < \gamma$,

$\mathcal{S} = \mathcal{S} \cup \omega_i$

else

$\mathcal{U} = \mathcal{U} \cup \omega_i$.

Set $\ell_{\omega_{i+1}} = \frac{\ell_{\omega_i}}{2}$.

Set $\epsilon = (\omega_i \in \mathcal{S} \ || \ \omega_i \in \mathcal{U})$

The sparse mode reduction algorithm will generate a set of modes that are guaranteed to be stable and achieve a desired robustness level γ, and "sparse", i.e. allowing modes such that at least one drug concentration is allowed to be zero, significantly reducing the number of the modes over which to apply the combination therapy algorithm. In the example described in Chapter 4, we synthesize controllers with

robustness level $\gamma = 14$. We start with $|\mathcal{P}|^m = 10000$ possible modes to search over, and reduce this number to 397 sparse modes.

3.5.3 A static state feedback combination therapy algorithm for nonlinear pharmacodynamics

As discussed in [42, 44] there are no known convex reformulations for the robust combination therapy problem due to the additional structure on L. As such we use the iterative approach developed in [44], as formalized in the previous section, based on the convex programs (3.12) and (3.13), to find a stabilizing controller, given a desired robustness level γ. For notation, let $Y' = \mathrm{P}_{Z'}(x, s)$ denote an optimization problem P in which we optimize over x and s leaving Z' fixed and with solution Y'. These optimization programs, taken from [83], are a synthesis variant of the conditions stated in Lemma 18.

Program 1. $\mathrm{P1}_{\ell,\omega}(x, s)$:

$$
\begin{aligned}
&\underset{x\in\mathbb{R}^n_+,s}{\text{minimize}}\ s \\
&\text{subject to} \\
&\quad A_\omega x + \Psi_\omega L x + \mathbf{1} \leq s \\
&\quad L = I \otimes \ell,\ \ell \in \omega \\
&\quad \mathbf{1}_n^T x \leq \gamma \\
&\quad s < 0, \qquad x \geq 0
\end{aligned}
\tag{3.22}
$$

Program 2. $\mathrm{P2}_{(x,\lambda_1,\lambda_2,\omega)}(\ell)$

$$
\begin{aligned}
&\underset{\ell\in\mathbb{R}^m_+}{\text{minimize}}\ \lambda_1\|\ell\|_1 + \lambda_2\|\ell\|_2 \\
&\text{subject to} \\
&\quad A_\omega x + \Psi_\omega L x + \mathbf{1} < 0 \\
&\quad L = I \otimes \ell,\ \ell \in \omega \\
&\quad \mathbf{1}_n^T x \leq \gamma
\end{aligned}
\tag{3.23}
$$

Algorithm 4 Scalable Combination Therapy For Nonlinear Pharmacodynamics

1. Set $\ell_0 = \ell_{\omega_{max}}$.
2. *Check if* $P1_{\ell_0,\omega}(x,s)$ *is feasible. If feasible*

   Set $(x', s') = \mathrm{P1}_{\ell_0,\omega}(x,s)$.
   Set $(\ell') = \mathrm{P2}_{(x',0,0,\omega)}(\ell)$.

   else, move to next mode and return to Step 1.
3. *Find* $(\lambda_1', \lambda_2', \ell_\omega)$ *for mode* ω:

   $\forall(\lambda_1, \lambda_2) \in \Lambda_1 \times \Lambda_2, \Lambda_1, \Lambda_2 \in \mathbb{R}^k_+$,
   Set $s = 1$.
   while $\neg(s' == s)$:
   Set $s = s'$.
   Set $(x', s') = \mathrm{P1}_{\ell',\omega}(x,s)$.
   Set $(\ell') = \mathrm{P2}_{(x',\lambda_1,\lambda_2,\omega)}(\ell)$.

Remark 22. Note the introduction of a slack variable s into Program 1, to help prevent immediate convergence to a local minimum. Minimizing s has the effect of maximizing the slack in the first constraint, allowing for more freedom in the design of the concentration vector ℓ in Program 2.

Chapter 4

Engineering Antibody Treatment Strategies to Control HIV

4.1 Introduction

A relatively recent discovery is that a minority of HIV-infected individuals can produce broadly neutralizing antibodies (bNAbs), that is, antibodies that inhibit infection by many strains of HIV [52]. These have been shown to inhibit infection by a broad range of viral isolates *in vitro* but also protect non-human primates against infection [5, 22, 52]. Passive transfer of human antibodies for the prevention and treatment of HIV-1 infection is increasingly being considered as an alternative to a conventional vaccine [14]. Only the most potent bNAbs are likely to be successful therapeutics, thus it is desirable from efficacy and monetary perspectives to deliver the best bNAb combinations. These bNAb combination treatment strategies aim to reproduce the success that resulted in combining antiretroviral drugs for the treatment of chronic HIV infection, and in addition have the potential of offering complete, long term viraemic control by enhancing host immunity [10, 69]. However, as the number of potential bNAbs grows, experimentally screening their combinations and concentrations for effectively controlling the evolution of HIV becomes increasingly infeasible. To address this, we require a scalable methodology that can take into account increasing amounts of HIV/bNAb resistance data, bNAb pharmacodynamic models and HIV mutational dynamics.

In Sections 3.4 and 3.5, we proposed a new scalable and computationally tractable algorithm that solves for optimal combinations and concentrations of bNAbs to neutralize virus in light of viral evolution while simultaneously allowing the designer to tailor treatment strategies in light of viral composition, maximum achievable doses, number of bNAbs used and ability to support pharmacodynamics/pharmacokinetic fluctuations, modeling and experimental error [43, 44]. Preliminary *in vitro* experimental data shows that our combination therapy algorithm predicts combinations and concentrations of bNAbs that can neutralize heterogeneous viral populations including those that include resistant mutants for one or more in the antibody mix (Figure 4.5).

The chapter is organized as follows: in Section 4.2, we demonstrate our ability to model control of the evolution to resistance of HIV in the presence of antibody therapy, through the application of the combination therapy algorithms developed in Chapter 3 as applied to experimental data derived from recent published studies [14, 22, 47]. We compare the $\mathcal{H}_\infty$ and the $\mathcal{L}_1$ combination therapy algorithms from Sections 3.3.3 and 3.4 with respect to their performance and robustness to biologically relevant uncertainty models and unmodeled dynamics. In Section 4.3, we discuss preliminary *in vitro* experimental methodology and results and show that the antibody treatment strategies synthesized with the nonlinear pharmacodynamics combination therapy algorithm described in Section 3.5.3 controls infection despite the presence of a mixed initial population of viruses, most of which are resistant to at least one antibody in the mix.

4.2 Mathematical Simulations

4.2.1 Model parameters

We consider a system of twenty eight HIV mutations with five potential antibodies to use in combination. Table 4.4 lists the mutations that evolved from monotherapy experiments with their corresponding half maximal inhibitory antibody concentra-

tion (IC50) in μg/ml, as measured in [47]. Antibodies 3BC176, PG16, 45-46G54W, PGT128 and 10-1074 are potential combination therapy candidates.

Infected cell replication rates. Although virus replication rates can vary considerably depending on the nature of the mutations a virus may undergo, we choose replication rates to be 0.5 $(\text{ml}\cdot\text{day})^{-1}$ for all mutants. We justify this selection by noting that escape mutants grew to be dominant mutants during selection experiments in [47] and assume that replication rate variability due to mutations were negligible.

Neutralization parameters. The fitness function associated with the neutralization of a virus i with respect to an antibody j is a Hill function

$$\psi_{ij} = \frac{\ell_j^n}{\ell_j^n + K_{ij}^n}, \tag{4.1}$$

where n is the Hill coefficient, ℓ_j is the concentration of a given antibody j,

$$K_{ij} = \frac{k_{on}}{k_{off}} = \frac{[x_i \ell_j]}{[x_i][\ell_j]} \tag{4.2}$$

is the association constant for the virus/antibody binding reaction $\ell_j + \mathrm{x_i} \xrightarrow{k_{on}} \ell_j \cdot \mathrm{x_i}$, and k_{on} and k_{off} are the on and off reaction rate constants. Note that the association constant represents the fraction bound of antibody/virus complexes in solution and that

$$K_{ij} = \frac{3 \cdot \mathrm{IC50}_{ij}}{3r_i + ln(2) - \mathrm{IC50}_{ij}}, \tag{4.3}$$

is found by solving Equation (4.6) for one virus/antibody pair for the duration $[t_0, t_f] = [0, 3]$. Results [22] show that Hill coefficients for CD4 binding site antibodies range from $n = 0.9$ to $n = 1.3$, and therefore we justify our simplification by setting the Hill coefficient $n = 1$. Our algorithm yields antibody concentrations near zero and this yields the linear approximation

$$\psi_{ij} = \frac{1}{K_{ij}} \ell_{ij}. \tag{4.4}$$

In addition, the antibodies we consider in our example do not target the same epitope,

in other words, do not bind competitively to the same sites on the virus, thereby reducing any coupling between antibody concentrations.

Mutation process. The mutation rate for HIV reverse transcriptase is $u = 3 \times 10^{-5}$ mutations/nucleotide base pair/replication cycle, and the HIV replication cycle is approximately 2.6 days. We approximate the rate of mutation for a particular amino acid mutation at a particular location to be $\frac{1}{n_a}u(1-u)^k = 1.443 \times 10^{-6}$ per replication cycle, where $k \approx 3000$ is the size of the genome in residues and $n_a = 19$ is the number of amino acids that can be mutated to. Our model supports forward point mutations and two point mutations. We do not consider back mutations, as the probability of mutation is negligible. Units of concentration in number of viruses/ml or number of antibodies/ml are used for states, and time is measured in days. The standard volume is 1 ml.

4.2.2 Controller synthesis

4.2.2.1 $\mathcal{L}_1$ controller synthesis

We synthesize a nominal stabilizing controller using Equation (3.7) comprised of an antibody pentamix (0.4687,0.7815, 0.6129, 0.6279, 0.8831) μg/ml of (3BC176, PG16, 45-46G54W, PGT128, 10-1074) using the convex program 6 and a robust controller using (3.8) that consisted of antibody trimix (0.6891,0.6712,1.0706) μg/ml of (3BC176, 4546-G54W, PGT128) using the $\mathcal{L}_1$ combination therapy algorithm. These were generated for the evolutionary dynamics with twenty eight HIV mutants listed in Table 4.4, Section 4.3.5.

Both antibody pentamix (stabilizing) and trimix (robustly stabilizing) controllers have similar gains and based on a cursory first glance, one might be believe these have comparable robustness properties. Indeed, for some simulations of the closed loop dynamics subjected to random time invariant perturbations in replication and neutralization dynamics, the nominal controller is stabilizing as seen in Figure 4.1. This is qualitatively consistent with the experimental results [47]. It was shown that with weekly injections of equal concentrations of the antibody pentamix described above,

viral loads remained below the limit of detection during an entire treatment course in mice. Moreover in [47], the presence of N160K-A281T-N332K, T162I-N280Y-N332K and T162I-N279K-N332K triple mutants were found after continuous antibody pentamix treatment in mice. Our simulation results show that these triple mutants dominate on the tenth day when the dynamics are destabilized by noise in the case of the nominally stabilizing controller. In particular, these triple mutants are resistant to all but one antibody 3BC176 in the pentamix, highlighting the importance of this antibody in the design of a robustly stabilizing controller. The robustly stabilizing antibody combination trimix synthesized by our algorithm addresses this particular dependence — a sparse combination is found that includes 3BC176 at a higher concentration and this trimix stabilizes HIV dynamics despite the presence of random perturbations in the dynamics.

In [47], an antibody trimix of equal concentrations of 3BC176, PG16 and 45-46G54W was suggested and experimentally shown to produce a decline in the initial viral load. However, a majority of mice in the experimental study had a viral rebound to pre-treatment levels, suggesting that in these cases, the virus had evolved mutations that were resistant to the trimix treatment. To compare the performance of our $\mathcal{L}_1$ synthesized controller with gains of (0.6891,0.6712,1.0706) μg/ml of (3BC176, 4546-G54W, PGT128) to the experimentally studied trimix, we chose equal concentrations of (3BC176, PG16, 45-46G54W), namely (1, 1, 1) μg/ml to mimic the experimentally derived trimix. We found that even though total antibody concentrations were larger in our version of the experimental trimix, the robustly stabilizing controller synthesized by the $\mathcal{L}_1$ algorithm nonetheless performed overall better; the closed loop induced 2-norm was $\|G\|_\infty = 0.2941$ and induced infinity norm was $\|G\|_{\infty-\text{ind}}= 0.6533$ for the $\mathcal{L}_1$ controller versus $\|G\|_\infty$=0.26433 and $\|G\|_{\infty-\text{ind}}$=0.74572 for the experimental trimix.

4.2.2.2 Performance comparisons between $\mathcal{H}_\infty$ and $\mathcal{L}_1$ controllers

We seek to provide a qualitative comparison between controllers synthesized using the scalable $\mathcal{L}_1$ and the $\mathcal{H}_\infty$ algorithms. To do this, we adapt the formulation in Section

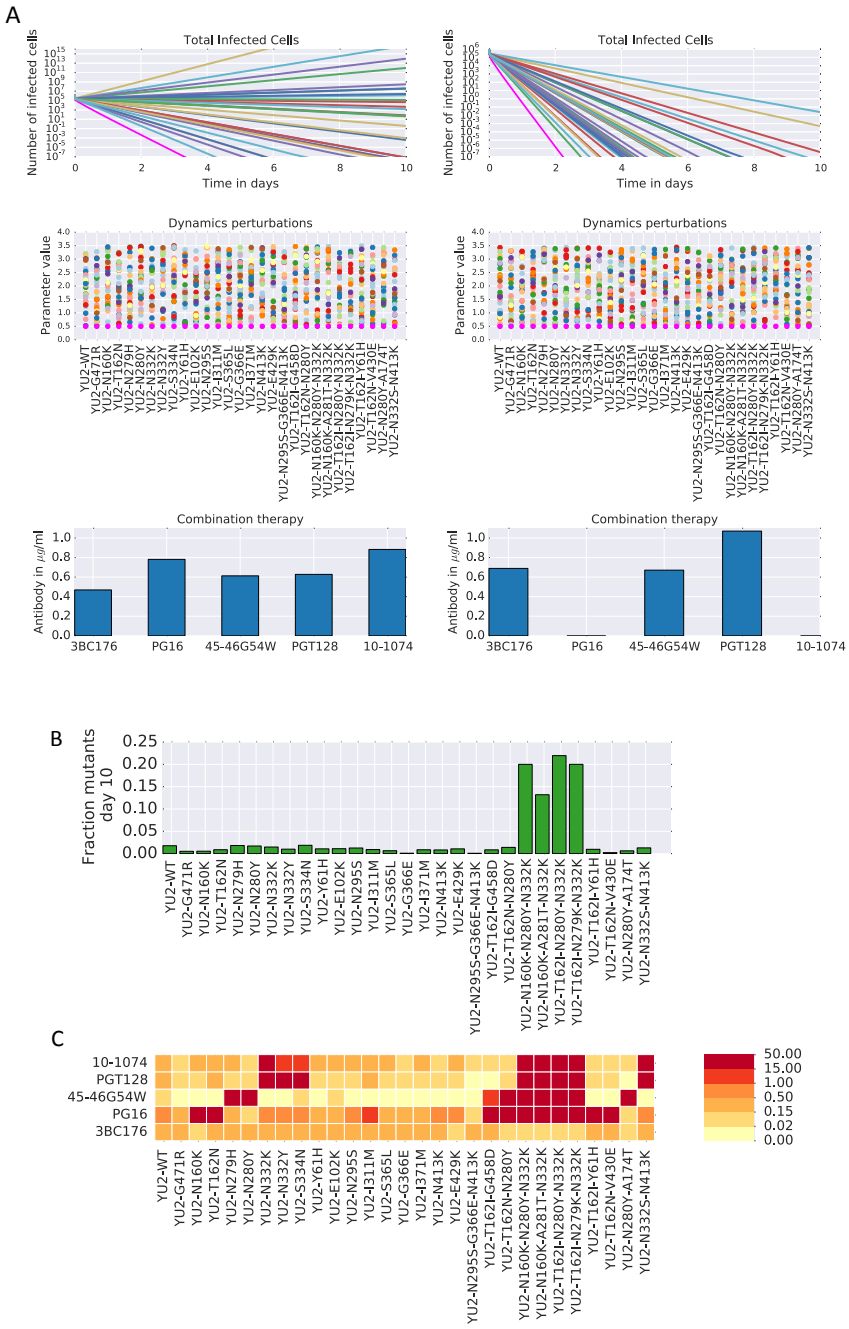

Figure 4.1: A. Sum of infected cell populations subject to random time invariant perturbations in the replication dynamics for 30 different simulations for (left) a stabilizing closed loop controller comprised of antibody pentamix (0.4687,0.7815, 0.6129, 0.6279, 0.8831) μg/ml of (3BC176, PG16, 45-46G54W, PGT128, 10-1074) synthesized using program (6) and (right) a robustly stabilizing closed loop controller comprised of antibody trimix (0.6891,0.6712,1.0706) μg/ml of (3BC176, 4546-G54W, PGT128) synthesized using the $\mathcal{L}_1$ combination therapy algorithm. B. (Left) Fraction of mutant infected cell populations remaining on day 10 for unstable simulations of HIV dynamics subject to noise, after the application of the antibody pentamix shown in A. (Right) Half maximal inhibitory concentrations of antibodies (IC50) with respect to HIV mutants, experimentally found in [47]. IC50 values are proportional to the degree of resistant for each mutant strain.

3.3.3 to include ℓ_1 and ℓ_2 regularization terms and solve the following non-convex problem using our iterative algorithm:

$$\begin{aligned}
&\text{minimize } \gamma + \lambda_1\|\ell\|_1 + \lambda_2\|\ell\|_2 \\
&\text{subject to} \\
&\begin{bmatrix} A_{cl}^T X + X A_{cl} + C^T C & X \\ X & -\gamma^2 I \end{bmatrix} \prec 0 \\
&\quad A_{cl} = (A - \Psi(\ell)) \\
&\quad C = \mathbf{1}^T \\
&\quad X \succ 0,\ X \text{ diagonal.}
\end{aligned} \tag{4.5}$$

We synthesized a nominal stabilizing controller using (3.7), a robust controller that minimizes the $\mathcal{H}_\infty$ closed loop norm using (3.8), and a robust controller using (3.23) that minimizes the $\mathcal{L}_1$ closed loop norm for the evolutionary dynamics of the eighteen HIV point mutants listed in Table 4.4. We found similar gains and robustness properties for both sparse and full controllers using either algorithm with the notable difference seen in computational time. Not surprisingly, the $\mathcal{L}_1$ algorithm has far superior performance, beating the runtime for the $\mathcal{H}_\infty$ synthesis algorithm by four orders of magnitude (Table 4.1).

We averaged thirty simulations of closed loop evolutionary dynamics subject to 5.5% random time invariant perturbations in the plant dynamics using both sparse and full support $\mathcal{H}_\infty$ and $\mathcal{L}_1$ controllers. The sparse controller found by the $\mathcal{L}_1$ algorithm performed better than the one found by $\mathcal{H}_\infty$ algorithm, whereas the situation was reversed for the respective synthesized full support controllers. As previously mentioned, the motivation for generating sparse controllers for combination therapy is that number of therapies commonly used in combination to treat a disease is often limited for clinical reasons. Therefore, the potential for the $\mathcal{L}_1$ algorithm to synthesize controllers that are not only sparse but more robustly stable than $\mathcal{H}_\infty$ algorithm with respect to parameter uncertainty and unmodeled dynamics is a desirable feature.

Figure 4.2 shows the relationship between gain sparsity and magnitude of (left) $\mathcal{H}_\infty$-induced norm and (right) ∞-induced norm in the synthesis of $\mathcal{H}_\infty$ and $\mathcal{L}_1$ con-

Controller	Gains μg/ml					$\|G\|_\infty$	$\|G\|_{\infty-\text{ind}}$	Time
	3BC176	PG16	45-46G54W	PGT128	10-1074			
Nominal Stabilizing	0.0125	0.0125	0.0125	0.0125	0.0125	2.7543	8.9786	1.53 s
Full $\mathcal{H}_\infty$	0.0550	0.0398	0.1587	0.1767	0.0629	0.1723	0.5988	>8 hours
Full $\mathcal{L}_1$	0.0424	0.0038	0.1125	0.2635	0.0548	0.1786	0.6722	35.803 s
Sparse $\mathcal{H}_\infty$	0	0	0.1485	0.1987	0	0.1917	0.6592	>8 hours
Sparse $\mathcal{L}_1$	0	0	0.1175	0.2971	0	0.1766	0.6616	35.803 s
Nominal Stabilizing (28 mutants)	0.4687	0.7815	0.6129	0.6279	0.8831	0.334	0.805	3.76 s
Sparse $\mathcal{L}_1$ (28 mutants)	0.6891	0	0.6712	1.0706	0	0.2941	0.6533	30.05 s

Table 4.1: Stabilizing gains found for nominal stabilizing controller, a robust controller using (3.8) that minimizes the $\mathcal{H}_\infty$ closed loop norm and a robust controller using (6) that minimizes the $\mathcal{L}_1$ norm of the closed loop system for evolutionary dynamics systems of the first eighteen HIV point mutants listed in Table 4.4, Appendix.

trollers. Although closed loop $\mathcal{H}_\infty$ norms remain constant with respect to sparsity for both types of controller synthesis, the closed loop ∞-induced norm decreases with sparsity for the $\mathcal{L}_1$ synthesized controller and increases with sparsity for the $\mathcal{H}_\infty$ synthesized controllers. This suggests that as expected, the $\mathcal{L}_1$ combination therapy algorithm finds better performing *sparse* controllers with respect to the ∞-induced norm than the $\mathcal{H}_\infty$ controller synthesis. Furthermore, performance guarantees with respect to the $\mathcal{H}_\infty$-induced norm are equivalent for both types of controller synthesis. As a result of computational limitations due to an SDP implementation of the $\mathcal{H}_\infty$ algorithm, we were were limited to synthesizing $\mathcal{H}_\infty$ and $\mathcal{L}_1$ controllers and comparing performance for a subset of eighteen mutants from Table 4.4.

These simulations demonstrate that although many stabilizing solutions to the combination therapy problem exist, the best ones are found when design parameters such as a sparsity, limits on the magnitude of gains, and robustness guarantees are simultaneously considered. Experimentally searching for these combinations is infeasible as the number of potential therapies and possible concentrations to consider is experimentally intractable. We propose to guide these experimental activities with our ability to design and synthesize combination therapy controllers. As such, one could generate a family of controllers based on "design specifications" tailored not only the (viral or cellular) composition of the disease, but to explore tradeoffs between number of therapies used (sparsity), therapy concentrations (magnitude of the gain) and ability to support pharmacokinetic fluctuations (robustness to perturbations) and

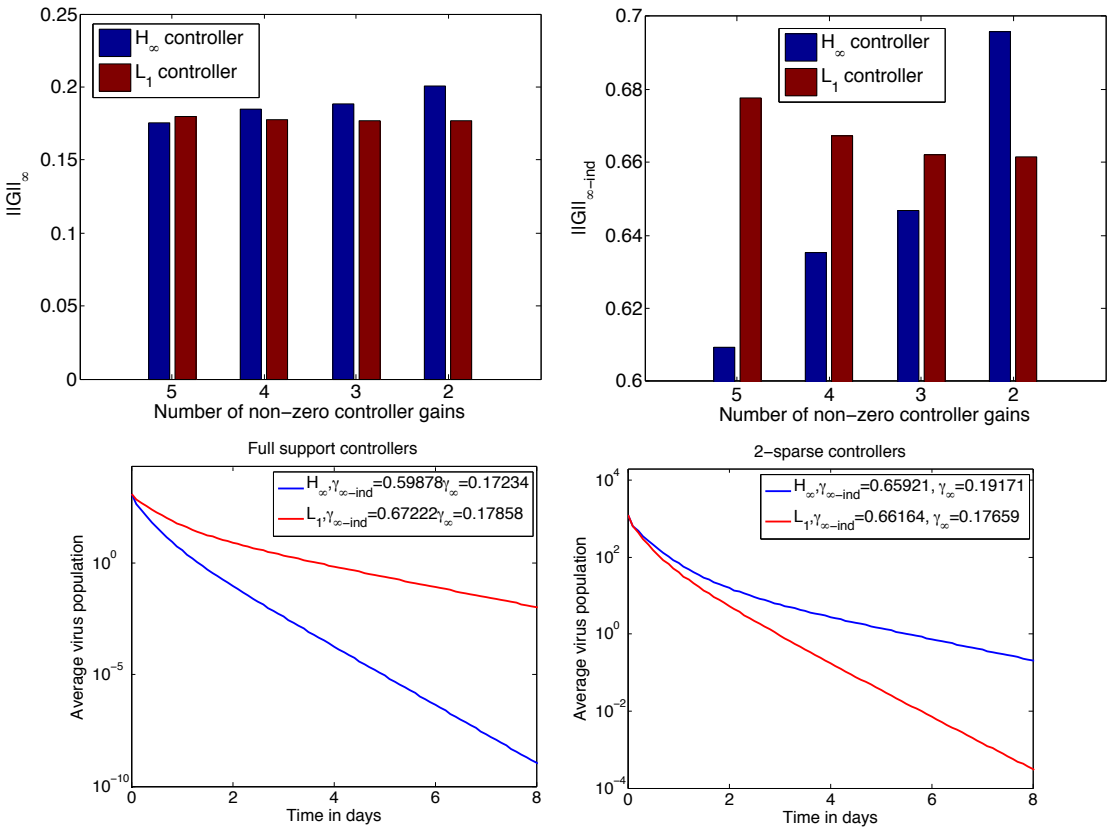

Figure 4.2: The graph depicts the average $\mathcal{H}_\infty$ (left) norm and $\infty - \text{ind}$ norm as a function of the sparsity of stabilizing controllers found using either the H_∞ or $\mathcal{L}_1$ combination therapy algorithms. The graphs depict the average of thirty simulations subject to random time invariant perturbations of 5.5 % in the plant dynamics found with the $\mathcal{H}_\infty$ and $\mathcal{L}_1$ combination therapy algorithms for evolutionary dynamics of the first 18 point mutants in Table 4.4, Appendix. (Left) Full support controllers synthesized with the pentamix of antibodies available: 3BC176, PG16, 45-46G54W, PGT128 and 10-1074 and (Right) Sparse controllers synthesized with only two antibodies 45-46G54W and PGT128.

subsequently verify these experimentally.

4.3 Preliminary Experimental Results

4.3.1 Abstract

Recent studies in HIV-1 immunotherapy have revealed that HIV can escape from the most potent broadly neutralizing antibody (bNAb) monotherapy and combination therapy. These data suggest a need to develop a methodology that directly addresses viral heterogeneity and that controls the evolutionary process leading to resistance in the design of effective bNAb treatment strategies. We developed a computational model and design tool that solves for candidate combinations and concentrations of bNAbs while allowing the exploration of trade offs in treatment design, such as limiting the number of bNAbs in the combination, dosage constraints and robustness to error. We demonstrate successful *in vitro* validation of our computationally predicted bNab combinations on heterogeneous viral populations comprised of resistant mutants. This study provides the first example of how combination anti-HIV bNAb treatment regimens can be rationally designed to maximize virus neutralization while minimizing the outgrowth of resistant populations.

4.3.2 Introduction

A challenge inherent to the treatment of chronic HIV infection is the risk that the virus will evolve to become resistant to treatment methods that comprise the standard of care. Recent pre-clinical and clinical data in the context of HIV immunotherapy have shown that HIV can escape from the most potent anti-HIV-1 broadly neutralizing antibody (bNAb) monotherapies and polytherapies [14, 37, 47, 95]. Heterogeneity in patient viral populations and the dynamics of evolutionary selection are issues that need to be explicitly addressed in the design of effective antibody combination regimens.

The challenge of designing treatment strategies that prevent resistance is one that

has been studied in both theoretical and experimental contexts. In the case of HIV and antiretroviral therapy, control theoretic methods have been proposed to delay the emergence of resistant mutant viruses, with the design of optimal drug scheduling and dosing strategies [25, 56, 108]. Other computational approaches integrate experimental data and modeling of antiretroviral pharmacodynamics [90], simulation and analysis of evolutionary dynamics models [31, 40, 87]. A computationally tractable method for the analysis and design of combination treatment strategies that is based on an experimental foundation and that allows for the principled exploration of design tradeoffs is missing.

Our proposed methodology for the rational design of combination therapy addresses these aforementioned issues, and in addition allows for the ability to quantifiably explore tradeoffs between number of drugs in therapy used in combination, their concentrations and the robustness of these treatment strategies to pharmacodynamic fluctuations and error. Our algorithm is based on the application of recent results in optimal control [82, 83] to an HIV evolutionary dynamics model and is constructed from experimentally derived antibody resistant phenotypes and their single antibody pharmacodynamics. This model assumes that individual antibodies when used in combination have independent, additive neutralization effects on single HIV mutants, which have been shown to hold for antibodies that bind at orthogonal binding sites [49], and is applicable in this study. This method represents a first step towards integrating principled engineering techniques with an experimentally-based mathematical model to design combination treatment strategies that offer predictive understanding of evolutionary dynamics and resistance of HIV with *in vitro* validation.

4.3.3 Results

Optimization of Combination Antibody Treatment Strategies

Our combination therapy algorithm identifies the smallest number of antibodies to combine and solves for the concentrations that robustly neutralizes HIV infection

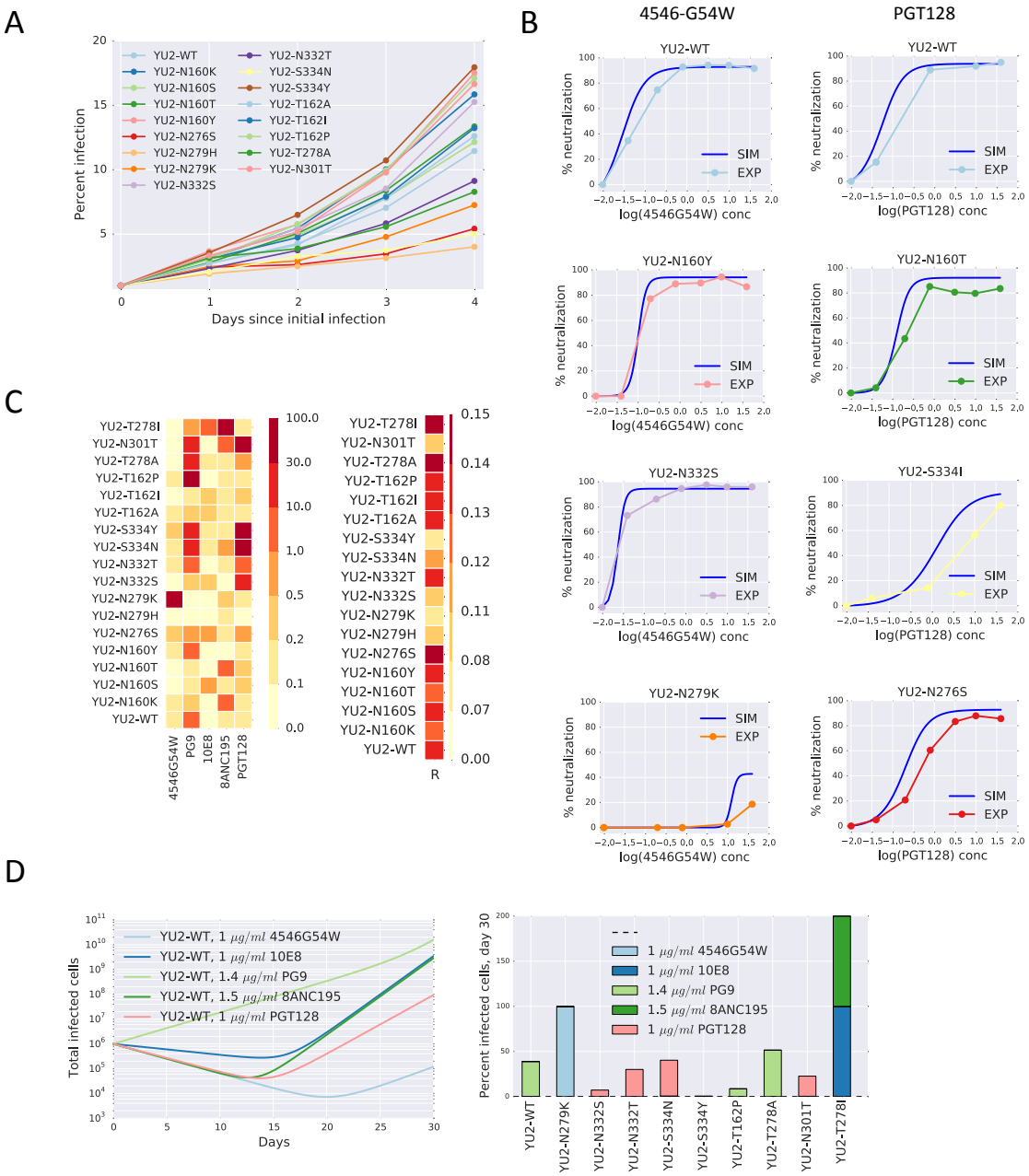

Figure 4.3: A. Experimentally derived infection timecourse for point mutants YU2-HIV. B. Experimentally derived neutralization curves for mutants of YU2-HIV with respect to antibodies 4546-G54W, PGT128. Results of evolutionary dynamics simulations run for 3 days, starting with 1 % infection of monoclonal YU2-WT, for different concentrations of antibodies 4546-G54W and PGT128, reproducing experimental conditions. C. Degree of resistance calculated as IC50×maximum neutralization. Replication rates for YU2-HIV mutants. D. Simulations of evolutionary dynamics for single antibody therapies. Percentage of virus subpopulations at day 30 for simulations shown.

subject to the evolutionary dynamics of mutation and selection. The goal is to design treatment strategies that stop the evolution to resistance or outgrowth of any population of resistant viruses, whether these viruses are pre-existing or occur as part of a selection process. Another feature of our methodology is that our combination therapies can quantifiably account for uncertainty and processes that are not a priori considered, such as experimental error, unmodeled resistance mechanisms [14] and changes in antibody pharmacokinetics [37]. Mathematical details of our evolutionary dynamics model and combination therapy algorithm can be found in Section 4.3.4.

In order to apply our algorithm, we first constructed our evolutionary dynamics model of replication, mutation and selection based on single antibody pharmacodynamics data derived for eighteen YU2-NL43 HIV-1 clade B Env mutants and five antibodies NIH4546G54W, PGT128, 8ANC195, PG9 and 10E8. We performed simulations and validated our experimentally derived evolutionary dynamics model both for consistency in replicating our own neutralization assays and also in predicting evolution to resistance (Figure 4.3B). For an initial population of monoclonal wild type YU2 exposed to a constant dose of a single antibody, we showed that mutants that evolved after 30 days from our simulations (Figure 4.3D) were consistent with previous evolution studies in humanized mice [14, 37, 47].

We then applied our design method to our experimentally derived evolutionary dynamics model to solve for the smallest set of antibody combinations and concentrations to neutralize infection. Our algorithm allows for the specification of hard constraints such as maximum allowable individual or total antibody concentrations, and soft constraints such as trade offs between robustness to uncertainty and dose minimization. Our methodology, which is based on the application of robust control theory, guarantees that if a combination therapy is found, it will be effective in neutralizing infection and controlling evolution to resistance for any mixture of viruses, given some knowledge of each of viruses in the mix. We specified that the total antibody concentration be less than 5 μg/ml and found two equivalently robust quadruple combinations that consisted of (1, 0, 1.5, 1.25, 1) μg/ml and (1, 1, 0, 1.5, 1) μg/ml of (45-46G54W, PGT128, 8ANC195, 10E8, PG9). Simulations of our evolu-

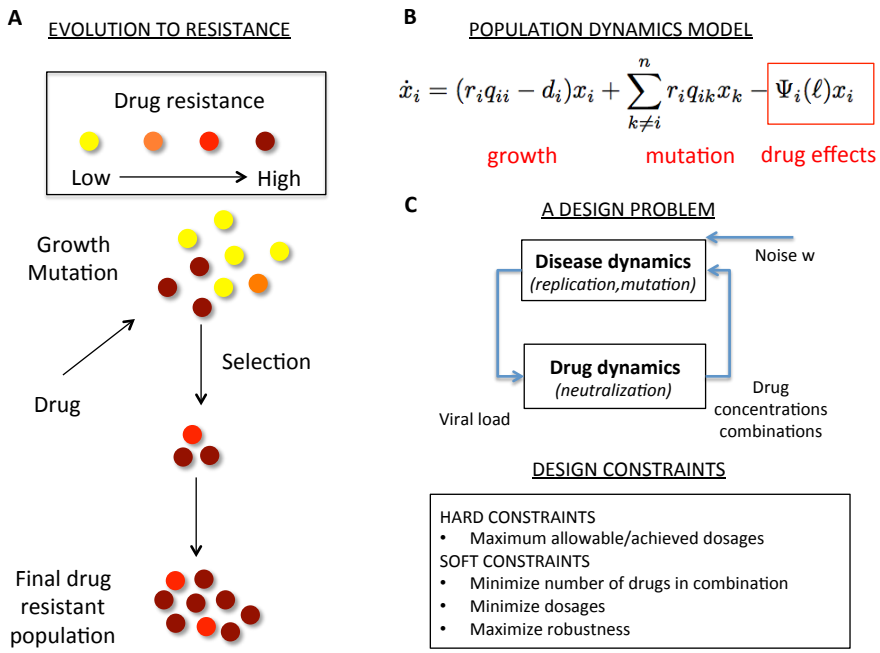

Figure 4.4: A. Schematic of growth, mutation and drug selection with respect to a population dynamics model. B. A deterministic differential equation model representing the population dynamics of replication, mutation of infectious units and their neutralization by antibodies. C. A schematic of the control system representing the combination therapy design problem, with examples of design constraints that can be specified. HIV replication and mutation dynamics are the unstable evolutionary process that we wish to control by finding effective combinations and concentrations of antibodies given particular treatment design constraints.

tionary dynamics model show that when our system is subjected to uncertainty, these combinations are still able to neutralize infection of a mixed population of resistant viruses (Figure 4.5). This particular example highlights that the best combinations are not necessarily the ones that include all antibodies, but ones that are designed with evolutionary dynamics and robustness in mind.

In Vitro Validation of Predicted Antibody Combinations.

To experimentally validate our predictions, we used an *in vitro* fluorescence based assay that exposed a mixture of wild type and mutant HIV infected cells to our predicted combination antibody therapies, where each mutant was resistant to at least one of the antibodies in the combination (Figure 4.5). Testing the combination antibody therapies predicted by our algorithm on heterogeneous viral populations allows us to verify that these regimens are agnostic to initial conditions consisting of resistant viral populations, as long as there is prior knowledge on their replication

fitness and single antibody response.

4.3.4 Materials and Methods

Mutagenesis, Virus Production and Cells. Site directed mutagenesis and assembly PCR were used to generate YU2-NL43 Env mutants. YU2-NL43 was modified using unique restriction sites EcoRI and XhoI. Inserts were generated by PCR using primers EcoRI-CF (5'-GCCAGCCAGAATTCTGCAACAACTGCTGTTTATCCAT TTCAG-3') and XhoI-CR-(5'-GCGTCGACCTCGAGATACTGCTCCCACCCCATC-3') and individual sense and antisense mutagenesis primers corresponding to YU2-NL43 mutants listed in the table below (Appendix). *Escherichia coli* One Shot STBL3 Chemically Competent cells (Life Technologies) were used to propagate proviral plasmids during a 16 hour incubation at 30° C. Stocks were prepared using a DNA Midi kit (Zymo Research). All gene constructs were verified by complete sequencing of gp160. Cell-free virus was produced by transfection of HEK293T cells with YU2-NL43 virus coding plasmid using BioT (Bioland Scientific). Viral supernatant was harvested at 48 h post transfection, filtered through a 0.5 μm filter and aliquots were stored at -80° C. Stock concentrations were quantified by p24 enzyme-linked immunosorbent assay (ELISA) (Cell Biolabs). The YU2 Env/NL43 plasmid was obtained from the Nussenzweig lab, Rockefeller University. The green fluorescent protein (GFP) reporter T-cell line GXR-CEM is previously described in [11] was obtained through AIDS Research and Reference Reagent Program, National Institute of Allergy and Infectious Diseases, National Institutes of Health.

Protein Expression and Purification. Antibodies were transiently expressed in HEK293T/17 cells or suspension HEK 293-6E cells (National Research Council Biotechnology Research Institute, Montréal, QC, Canada) using 25-kDa linear polyethylenimine (Polysciences) for transfection. Supernatants were passed over MabSelect SuRe protein A resin (GE Healthcare) or Protein G Sepharose 4 Fast Flow (GE Healthcare) and eluted by using pH 3.0 citrate or glycine buffer, and then immediately neutralized. Antibodies were further purified by size exclusion chromatography

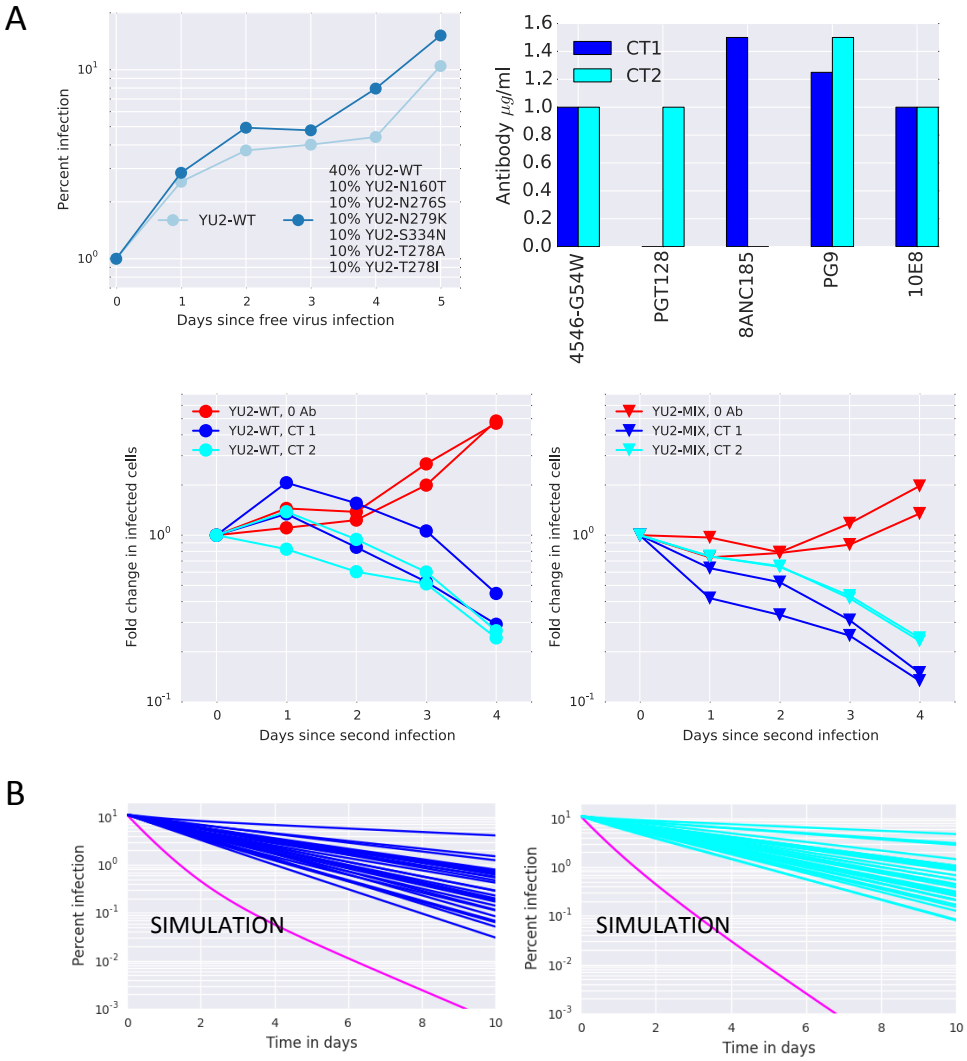

Figure 4.5: A. Combination therapy experiment showing the timecourse of infection after administering two different antibody combinations that were solved with the combination therapy algorithm, for two different initial conditions: YU2-WT and a mix of YU2-WT and of YU2-(N160T,N276S, N279K, S334N, T278A, T278I) each one resistant to one or more of (45-46G54W, PGT128, 10E8, 8ANC195, PG9). Infection timecourse shown for both initial conditions. B. Thirty simulations of evolutionary dynamics subject to random time invariant perturbations of 8% in the system dynamics, given robust combination therapy 1 (CT1) (Blue), a robust combination therapy 2 (CT2). The magenta line corresponds to the evolutionary dynamics of the unperturbed system.

using a Superdex 200 or 75 10/300 GL column.

In Vitro Replication and Neutralization Assays. To initiate infection for both replication and neutralization assays, GXR-CEM cells at 4×10^5 cells/ml were infected with 200 ng YU2-NL43 (HIV) virus stock in the absence of antibody and incubated for three days at 37° C. Two days after infection, uninfected GXR-CEM cells at 2×10^5 were pre-treated with 0, 0.02,0.08, 0.4,1.6, 6.4, 20, 80 μg/ml of single PGT128, NIH45-46^{G54W}, 8ANC195, PG9 and 10E8 antibodies. Three days after initial infection with cell free virus, the infected GXR-CEM were washed and added to a final concentration of 1 % GFP-expressing donor cells to uninfected pre-treated GXR-CEM cells. For the neutralization assay, a constant concentration of antibody was maintained for each sample for three days following secondary infection. Infection was determined by measuring GFP reporter gene expression in the absence of antibodies measured daily for four days following secondary infection. Neutralization was determined by measuring the reduction in GFP reporter gene expression in the presence of antibodies PGT128, NIH4546G54W, 8ANC195, PG9 and 10E8 daily for three days. Nonlinear regression analysis was used to calculate the concentrations at which half-maximal inhibition was observed (IC50 values) by fitting the observed neutralization to the expression $\gamma\ell^n/(\ell^n + \text{IC50}^n)$, where γ is the maximum achievable neutralization, ℓ is the antibody concentration and n is the Hill coefficient.

In Vitro Combination Therapy Assays. To initiate infection for the mixed mutant initial condition, GXR-CEM cells at 4×10^5 cells/ml were infected with (80, 20, 20, 20, 20, 20) ng with each of YU2-NL43-(WT, S334N, N279K, N276S, T278I,N160T,T278A) virus stock in the absence of antibody and incubated for three days. Three days after initial infection with cell free virus, the infected GXR-CEM were washed and added to a final concentration of 1 % GFP-expressing donor cells to uninfected pre-treated GXR-CEM cells. On each day after infected donor addition, cell aggregates were broken up by gentle repeated pipetting and measured by FACS for four days. Infection conditions were calibrated so that the number of uninfected target cells would not be limiting and infection would not interfere with proliferation of uninfected cells.

Flow cytometry and Dynamical System Parameter Estimation. Flow cytometry data was collected with a MACSQuant flow cytometer and transformed to a python format using the FlowCytometryTools python package (Gore lab, MIT). An automated gating and analysis software module written in python was developed to process large sets of flow cytometry data. All parameter estimation was done using the python scipy.optimize module. Dynamical system replication rates for YU2-NL43 mutants were calculated by fitting an exponential growth function $x(t) = x(0)e^{rt}$. Neutralization rates were calculated by fitting a Hill equation $\gamma \frac{\ell^n}{\ell^n + \text{IC50}^n}$, where ℓ is antibody concentration, γ is the maximum neutralization, n is the Hill coefficient, and IC50 is half maximal inhibitory concentration curves. Mutation rates were calculated using the mutation rate for HIV reverse transcriptase $\mu = 3 \times 10^{-5}$ mutations/base pair/replication cycle [] for the specific YU2-NL43 mutants that were constructed.

Model of HIV evolutionary dynamics. We use a deterministic model of HIV evolutionary dynamics to model replication, mutation and neutralization:

$$\dot{x}_i = r_i q_{ii} x_i + \sum_{k \neq i}^{n} r_i q_{ik} x_k - \sum_{j=1}^{m} \gamma_{ij} \frac{\ell_j^{n_{ij}}}{\ell_j^{n_{ij}} + \text{IC50}^{n_{ij}}} x_i, \tag{4.6}$$

where $x_i \in \mathbb{R}_+$ is the concentration of infectious mutant i, $\ell_j \in \mathbb{R}$ is the antibody concentration in μg/ml (assumed to remain at constant concentrations throughout), r_i is the replication rate of mutant i, and q_{ik} is the probability that mutant k mutates to mutant i (note that q_{ii} is the probability of no mutation occurring). The rates r_i can be viewed as the infection rates of mutant i without the effect of the antibody. γ_{ij} and n_{ij} are the maximum achievable neutralization and the Hill coefficient for the neutralization of mutant i by antibody j. This model assumes that individual antibodies when used in combination have independent, additive neutralization effects. This does not overly constrain our problem as we generally consider antibodies that bind at orthogonal binding sites which have been shown to have independent, additive effects when used in combination.

Combination Therapy Algorithm. We applied the combination therapy algorithm for nonlinear pharmacodynamics described in Section 3.5.3, using the model of

evolutionary dynamics described above, with parameters listed in Tables 4.2 and 4.3, Appendix. We first generated a set of 10000 pharmacodynamic modes that represent piecewise linearizations of the neutralization curves. Using our sparse mode reduction algorithm (3), we reduced this number to 397 modes and used these as input to the combination therapy algorithm. Several design constraints were defined prior to running the combination therapy algorithm: 1) the total antibody concentration be less than 5 μg/ml, 2) the combination therapy must be robust to perturbations. Given these constraints, the algorithm solved for two quadruple antibody combinations with equivalent robustness levels.

4.3.5 Appendix

	R		45-46G54W			PGT128	
		max	IC50	n	max	IC50	n
YU2-WT	0.646	0.94234	0.06636	1.3945	0.94924	0.12469	1.4692
YU2-N160K	0.619	0.91966	0.14658	2.9826	0.98216	0.12977	0.65651
YU2-N160S	0.704	0.94892	0.055377	0.99516	0.90636	0.29987	1.4888
YU2-N160T	0.637	1	0.030035	4.3297	0.85211	0.19524	2.5482
YU2-N160Y	0.66	0.94412	0.13862	4.0697	0.95484	0.14137	3.4386
YU2-N276S	0.724	0.74572	0.27784	1.4011	0.87891	0.45284	1.4061
YU2-N279H	0.426	0.8932	0.027408	7.9821	1	0.020018	10.947
YU2-N279K	0.362	0.18648	13.084	6.4848	0.88866	0.079372	1.0901
YU2-N332S	0.503	0.97656	0.031375	4.5121	0.73823	19.925	10.703
YU2-N332T	0.689	0.94674	0.13681	2.7212	0.1366	0.18446	4.7561
YU2-S334N	0.562	0.90953	0.13689	2.1111	0	7.0653	0.93399
YU2-S334Y	0.42	1	0.25627	0.84276	0.37354	19.99	9.7835
YU2-T162A	0.734	1	0.12631	1.7346	0.97392	0.1324	1.4749
YU2-T162I	0.643	0.96232	0.032941	3.1854	0.96036	0.20933	1.618
YU2-T162P	0.655	0.91681	0.13396	2.2804	1	0.16072	1.1002
YU2-T278A	0.72	0.96287	0.030657	5.2389	0.94765	0.54778	1.9417
YU2-N301T	0.542	0.91602	0.037664	5.6922	0	1.2506	0.82346
YU2-T278I	0.713	0.95487	0.052562	1.558	0.96781	0.17467	1.0706

Table 4.2: Replication rates per day and neutralization curve parameters for the indicated antibodies on YU2-NL43 mutant viruses.

		10E8			8ANC195			PG9	
		max	IC50	n	max	IC50	n	max	IC50
YU2-WT	0.94486	0.060457	0.47707	0.9347	0.12696	1.7026	0.45589	0.49713	0.70523
YU2-N160K	0.96615	0.022256	0.559	0.49399	1.5986	9.8458	0.3395	0.091707	0.30884
YU2-N160S	0.84544	0.60074	0.37193	0.98259	0.074471	0.64653	0.59807	0.074839	0.41778
YU2-N160T	0.92565	0.013653	0.77443	0.91221	1.5885	11.62	0.59807	0.074839	0.41778
YU2-N160Y	0.95313	0.0073578	1.0627	1	0.051382	0.31669	0.68767	1.4407	1.375
YU2-N276S	0.94657	0.70223	0.51266	0.80248	0.14602	7.3062	0.70541	0.67781	0.86789
YU2-N279H	0.97946	0.0057117	2.3761	0.97464	0.14855	7.9667	0.9263	0.057702	0.92412
YU2-N279K	0.90362	0.011933	0.69676	0.81632	0.22242	0.60999	0.9263	0.057702	0.92412
YU2-N332S	0.95397	0.22654	0.49865	0.94235	0.03849	1.1722	0.76037	0.30023	0.57322
YU2-N332T	0.96289	0.050848	0.43029	0.90555	0.16991	1.0793	0.71449	4.2677	0.9406
YU2-S334N	0.85623	0.10818	1.0919	0.85539	0.65643	1.9121	0.59587	11.837	7.5342
YU2-S334Y	0.87005	0.065392	1.0087	0.84908	0.12062	0.92592	0.6376	12.24	1.4706
YU2-T162A	0.95998	0.24036	0.51997	0.89124	0.15116	0.82269	0.58197	0.25439	0.55699
YU2-T162I	0.93713	0.24137	0.58461	0.82959	0.11562	0.79189	0.45277	0.075214	0.4691
YU2-T162P	0.9602	0.013527	0.54256	0.83843	0.11323	0.88195	0.55945	19.838	11.047
YU2-T278A	0.89614	0.16974	0.43169	0.64413	0.074199	0.50969	0.84999	9.6801	7.6016
YU2-N301T	0.90295	0.025051	0.62866	0.75917	0.86499	7.1084	0.69971	20.099	11.154
YU2-T278I	0.83407	6.4431	4.9855	0	2.6138	0.049654	0.86148	0.46376	0.42271

Table 4.3: Neutralization curve parameters for the indicated antibodies on YU2-NL43 mutant viruses.

Antibody associated escape mutants	Mutation	3BC176 IC50 μg/ml	PG16 IC50 μg/ml	45-46G54W IC50 μg/ml	PGT128 IC50 μg/ml	10-1074 IC50 μg/ml
	WT	0.319	0.612	0.024	0.169	0.312
3BC176	G471R	0.159	0.154	0.008	0.02	0.091
PG16	N160K	0.145	50	0.007	0.086	0.155
	T162N	0.154	50	0.013	0.166	0.175
45-46G54W	N279H	0.209	0.294	50	0.064	0.177
	N280Y	0.276	0.145	50	0.031	0.126
PGT128 or 10-1074	N332K	0.232	0.988	0.017	50	50
	N332Y	0.269	0.632	0.01	50	13.596
	S334N	0.218	0.615	0.02	50	7.308
Passenger mutations	Y61H	0.243	0.285	0.015	0.098	0.26
	E102K	0.173	0.341	0.023	0.11	0.207
	N295S	0.347	0.5	0.017	0.145	0.159
	I311M	0.23	2.67	0.013	0.248	0.253
	S365L	0.26	0.273	0.009	0.045	0.153
	G366E	0.187	0.167	0.001	0.021	0.074
	I371M	0.2	0.303	0.013	0.064	0.164
	N413K	0.188	0.557	0.014	0.032	0.109
	E429K	0.146	0.503	0.017	0.082	0.167
	N295S-G366E-N413K	0.222	0.131	0.001	0.012	0.021
tri-mix	T162I-G458D	0.275	50	14.33	0.012	0.047
	T162N-N280Y	0.138	50	50	0.027	0.079
penta-mix	N160K-N280Y-N332K	0.146	50	50	50	50
	N160K-A281T-N332K	0.1	50	50	50	50
	T162I-N280Y-N332K	0.13	50	50	50	50
	T162I-N279K-N332K	0.149	50	50	50	50
Signature + Passenger	T162I-Y61H	0.156	50	0.014	0.088	0.115
	T162N-V430E	0.167	50	0.003	0.037	0.106
	N280Y-A174T	0.064	0.138	50	0.01	0.021
	N332S-N413K	0.181	0.526	0.017	50	50
Estimated Mutations	N160K-N280Y	0.276	50	50	0.086	0.155
	N160K-N332K	0.232	50	0.017	50	50
	N280Y-N332K	0.276	0.988	50	50	50
	N295S-G366E	0.347	0.5	0.017	0.145	0.159
	N295S-N413K	0.347	0.557	0.017	0.145	0.159
	G366E-N413K	0.188	0.557	0.014	0.032	0.109

Table 4.4: IC50 values for the indicated antibodies on YU2 mutant viruses found in continuous antibody mono therapy experiments conducted by the Nussenzweig lab at Rockefeller University [47]. The trimix of antibodies is : 3BC176,PG16,45-46G54W, the penta-mix is 3BC176, PG16, 45-46G54W , PGT128 and 10-1074. Estimated two point mutations represent intermediary mutations needed for our model but not included in experimental results in [47]. The IC50 values were taken to be the maximum IC50 of both mutations.

Chapter 5

Summary and Future Directions

This thesis describes the development, analysis and implementation of novel mathematical models and control theoretic algorithms for the prediction and control of evolutionary dynamics disease models. In particular, we apply these methods to study the evolution of HIV resistance in the presence of broadly neutralizing antibody therapy and to synthesize effective combination treatment strategies for effective control. We show that our models and techniques are validated with our own experimental studies as well as in agreement with previous studies. The contributions of this thesis are summarized below.

Evolutionary Dynamics on Computationally Derived Fitness Landscapes

We develop two computational models to reason about HIV resistance to antibody therapy. Our first method is an extension of the least absolute shrinkage and selection operator (LASSO) that serves to identify potential escape mutations from neutralization data. Our model is applied to HIV and antibody neutralization data and uncovers key residue locations and substitutions that are at the basis of neutralization, that are verified by previous or our own experimental studies.

Our second computational model allows us to reason about the dynamics of evolution on computationally generated fitness landscapes. It is developed from biophysical first principles that link differences in binding energies with changes in replication

and the effect of selective pressure. We apply our approach to an HIV evolutionary dynamics model that incorporates infection and antibody neutralization dynamics, a mutation process, and a method that uses energy minimization calculations on structural information to quantify fitness differences between sensitive and resistant strains. We show that there is agreement between our model and previous murine studies of HIV evolution in the presence of antibody therapy. A useful extension is to include a mutational model that captures the dynamics of two mutations or more to provide further predictive power.

Robust Control of Evolutionary Dynamics

We present three algorithms for the principled design of targeted combination drug treatment strategies that explicitly account for the evolutionary dynamics of a generic disease model, where the drugs under consideration are non-interacting and exhibit independent additive effects. These algorithms allow the designer to quantifiably explore tradeoffs between number of therapies used (controller sparsity), therapy concentrations (magnitude of the gain) and ability to support pharmacokinetic fluctuations (robustness to perturbations).

Our first algorithm proposes a general iterative method that uses an $\mathcal{H}_\infty$ robust control approach to design targeted combination therapy concentrations and is effective in generating robustly stabilizing controllers. Our second algorithm addresses large scale systems concerns lacking in the first algorithm, presenting a scalable solution to the combination therapy problem by reformulating it as a second order cone program (SOCP), with robustness guarantees addressed by minimization of the induced $\mathcal{L}_1$ norm. Our third algorithm solves the combination therapy problem subject to the same design constraints (sparsity of the drug combination, maximum dosage and robustness constraints) formulated as an SOCP while addressing the nonlinear dynamics of individual drugs and of their combinations.

Engineering Antibody Treatment Strategies to Control HIV

We demonstrate our ability to control the evolution to resistance of HIV in the presence of antibody therapy, through the application of the combination therapy algorithms developed in Chapter 3 as applied to experimental data derived from recent published studies [14, 22, 47]. Specifically, we synthesize combination treatments and compare the respective $\mathcal{H}_\infty$ and the $\mathcal{L}_1$ combination therapy algorithms with respect to their performance and robustness to biologically relevant uncertainty models and unmodeled dynamics.

We also develop an *in vitro* experimental methodology and demonstrate that the antibody treatment strategies synthesized with the nonlinear pharmacodynamics combination therapy algorithm described in Section 3.5.3 controls infection despite the presence of a mixed initial population of viruses, most of which are resistant to at least one antibody in the mix. Additional work is in progress to further experimentally validate other combination strategies synthesized by our algorithm.

Implications

With the expansion of molecular data associated with human disease and the rapid discovery of new targeted therapies, the application of individualized medicine is becoming a more tangible prospect. The ability to do this not only requires that health care providers have access to large sets of disease-related data but that they are also equipped to use it to engineer effective treatments for their patients.

The research in this thesis centers around the idea that precision medicine is ultimately the combination of multiple disciplines: the understanding of biomolecular foundations and dynamics of disease, the consolidation and analysis of large sets of molecular and clinical data, and the ability to incorporate these complementary views on disease to engineer treatment strategies tailored to a patient's disease population, pharmacokinetic limitations all while being robust to evolution to resistance. Our goal was to provide a theoretical and computational foundation for a set of predictive tools to help reason about the nature of genetic alterations that lead to resistance, the

disease dynamics leading to escape from targeted therapy and to synthesize treatment strategies to control evolution of disease. The methods presented here are applied to the prediction and control of HIV resistance in the presence of antibody therapy, but are currently being extended to address tumor heterogeneity and control evolution in cancer.

Bibliography

[1] B. Al-Lazikani, U. Banerji, and P. Workman. Combinatorial drug therapy for cancer in the post-genomic era. *Nature Biotechnology*, 30(7):679–692, 2012.

[2] M. Alamir. Robust feedback design for combined therapy of cancer. *Optimal Control Applications and Methods*, 2012.

[3] M. Alamir and S. Chareyron. State-constrained optimal control applied to cell-cycle-specific cancer chemotherapy. *Optimal Control Applications and Methods*, 28(3):175–190, 2007.

[4] P. A. Ascierto, D. Schadendorf, C. Berking, S. S. Agarwala, C. M. van Herpen, P. Queirolo, C. U. Blank, A. Hauschild, J. T. Beck, A. St-Pierre, et al. Mek162 for patients with advanced melanoma harbouring nras or val600 braf mutations: a non-randomised, open-label phase 2 study. *The Lancet Oncology*, 14(3):249–256, 2013.

[5] A. B. Balazs, J. Chen, C. M. Hong, D. S. Rao, L. Yang, and D. Baltimore. Antibody-based protection against HIV infection by vectored immunoprophylaxis. *Nature*, 481(7379):81–84, 2011.

[6] G. Barbato, E. Bianchi, P. Ingallinella, W. H. Hurni, M. D. Miller, G. Ciliberto, R. Cortese, R. Bazzo, J. W. Shiver, and A. Pessi. Structural analysis of the epitope of the anti-HIV antibody 2F5 sheds light into its mechanism of neutralization and HIV fusion. *Journal of Molecular Biology*, 330(5):1101–1115, 2003.

[7] C. M. Blakely, E. Pazarentzos, V. Olivas, S. Asthana, J. J. Yan, I. Tan, G. Hrustanovic, E. Chan, L. Lin, D. S. Neel, et al. Nf-κb-activating complex engaged in response to egfr oncogene inhibition drives tumor cell survival and residual disease in lung cancer. *Cell Reports*, 11(1):98–110, 2015.

[8] F. Blanchini, P. Colaneri, and R. H. Middleton. A convexity result for the optimal control of a class of positive nonlinear systems. In *19th IFAC World Congress*, 2014.

[9] C. A. Blish, Z. Jalalian-Lechak, S. Rainwater, M.-A. Nguyen, O. C. Dogan, and J. Overbaugh. Cross-subtype neutralization sensitivity despite monoclonal antibody resistance among early subtype A, C, and D envelope variants of human immunodeficiency virus type 1. *Journal of Virology*, 83(15):7783–7788, 2009.

[10] S. Bournazos, F. Klein, J. Pietzsch, M. S. Seaman, M. C. Nussenzweig, and J. V. Ravetch. Broadly neutralizing anti-HIV-1 antibodies require Fc effector functions for in vivo activity. *Cell*, 158(6):1243–1253, 2014.

[11] M. A. Brockman, G. O. Tanzi, B. D. Walker, and T. M. Allen. Use of a novel GFP reporter cell line to examine replication capacity of CXCR4-and CCR5-tropic HIV-1 by flow cytometry. *Journal of Virological Methods*, 131(2):134–142, 2006.

[12] S. Bryson, J.-P. Julien, R. C. Hynes, and E. F. Pai. Crystallographic definition of the epitope promiscuity of the broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2F5: vaccine design implications. *Journal of Virology*, 83(22):11862–11875, 2009.

[13] E. Candès and B. Recht. Exact matrix completion via convex optimization. *Commun. ACM*, 55(6):111–119, June 2012.

[14] M. Caskey, F. Klein, J. C. Lorenzi, M. S. Seaman, A. P. West Jr, N. Buckley, G. Kremer, L. Nogueira, M. Braunschweig, J. F. Scheid, et al. Viraemia sup-

pressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117. *Nature*, 2015.

[15] V. Chandrasekaran, B. Recht, P. Parrilo, and A. Willsky. The convex geometry of linear inverse problems. *Foundations of Computational Mathematics*, 12:805–849, 2012.

[16] S. Chareyron and M. Alamir. Model-free feedback design for a mixed cancer therapy. *Biotechnology Progress*, 25(3):690–700, 2009.

[17] E. Chertova, J. W. Bess Jr, B. J. Crise, R. C. Sowder II, T. M. Schaden, J. M. Hilburn, J. A. Hoxie, R. E. Benveniste, J. D. Lifson, L. E. Henderson, et al. Envelope glycoprotein incorporation, not shedding of surface envelope glycoprotein (gp120/SU), is the primary determinant of su content of purified human immunodeficiency virus type 1 and simian immunodeficiency virus. *Journal of Virology*, 76(11):5315–5325, 2002.

[18] A. T. C. Collaboration et al. Life expectancy of individuals on combination antiretroviral therapy in high-income countries: a collaborative analysis of 14 cohort studies. *The Lancet*, 372(9635):293–299, 2008.

[19] M. Dahleh. *l1 Robust Control: Theory, Computation and Design*. CRC Press, Boca Raton, FL, 1995.

[20] C. Darwin. On the origin of the species by natural selection. 1859.

[21] N. Dhingra, F. Lin, M. Fardad, and M. R. Jovanovic. On identifying sparse representations of consensus networks. In *3rd IFAC Workshop on Distributed Estimation and Control in Networked Systems, Santa Barbara, CA*, pages 305–310, 2012.

[22] R. Diskin, J. F. Scheid, P. M. Marcovecchio, A. P. West, F. Klein, H. Gao, P. N. Gnanapragasam, A. Abadir, M. S. Seaman, M. C. Nussenzweig, et al. Increasing the potency and breadth of an HIV antibody by using structure-based rational design. *Science*, 334(6060):1289–1293, 2011.

[23] R. C. Doebele, A. B. Pilling, D. L. Aisner, T. G. Kutateladze, A. T. Le, A. J. Weickhardt, K. L. Kondo, D. J. Linderman, L. E. Heasley, W. A. Franklin, et al. Mechanisms of resistance to crizotinib in patients with ALK gene rearranged non-small cell lung cancer. *Clinical Cancer Research*, 18(5):1472–1482, 2012.

[24] D. L. Donoho. Compressed sensing. *IEEE Trans. Inform. Theory*, 52:1289–1306, 2006.

[25] R. M. E. HernandezVargas, P. Colaneri and F. Blanchini. Discrete-time control for switched positive systems with application to mitigating viral escape. *International Journal on Robust and Nonlinear Control*, 21:1093–1111, 2011.

[26] M. Eigen, J. McCaskill, and P. Schuster. The molecular quasi-species. *Adv. Chem. Phys*, 75:149–263, 1989.

[27] J. A. Engelman and P. A. Janne. Mechanisms of acquired resistance to epidermal growth factor receptor tyrosine kinase inhibitors in non–small cell lung cancer. *Clinical Cancer Research*, 14(10):2895–2899, 2008.

[28] M. Fardad, F. Lin, and M. R. Jovanovic. Design of optimal sparse interconnection graphs for synchronization of oscillator networks. *IEEE Transactions on Automatic Control*, 59(9):2457–2462, 2014.

[29] M. A. Fischl, D. D. Richman, M. H. Grieco, M. S. Gottlieb, P. A. Volberding, O. L. Laskin, J. M. Leedom, J. E. Groopman, D. Mildvan, R. T. Schooley, et al. The efficacy of azidothymidine (azt) in the treatment of patients with aids and aids-related complex. *New England Journal of Medicine*, 317(4):185–191, 1987.

[30] R. A. Fisher. *The genetical theory of natural selection: a complete variorum edition*. Oxford University Press, 1930.

[31] F. Fu, M. A. Nowak, and S. Bonhoeffer. Spatial heterogeneity in drug concentrations can facilitate the emergence of resistance to cancer therapy. *PLoS Computational Biology*, 11(3):e1004142–e1004142, 2015.

[32] I. S. Georgiev, N. A. Doria-Rose, T. Zhou, Y. Do Kwon, R. P. Staupe, S. Moquin, G.-Y. Chuang, M. K. Louder, S. D. Schmidt, H. R. Altae-Tran, et al. Delineating antibody recognition in polyclonal sera from patterns of HIV-1 isolate neutralization. *Science*, 340(6133):751–756, 2013.

[33] J. B. S. Haldane. The origin of life. *Rationalist Annual*, 148:3–10, 1929.

[34] J. S. Haldane. The rate of mutation of human genes. *Hereditas*, 35(S1):267–273, 1949.

[35] A. Hochhaus, S. Kreil, A. Corbin, P. La Rosee, M. Muller, T. Lahaye, B. Hanfstein, C. Schoch, N. Cross, U. Berger, et al. Molecular and chromosomal mechanisms of resistance to imatinib (STI571) therapy. *Leukemia*, 16(11):2190–2196, 2002.

[36] S. Hoelder, P. A. Clarke, and P. Workman. Discovery of small molecule cancer drugs: successes, challenges and opportunities. *Molecular Oncology*, 6(2):155–176, 2012.

[37] J. A. Horwitz, A. Halper-Stromberg, H. Mouquet, A. D. Gitlin, A. Tretiakova, T. R. Eisenreich, M. Malbec, S. Gravemann, E. Billerbeck, M. Dorner, et al. HIV-1 suppression and durable control by combining single broadly neutralizing antibodies and antiretroviral drugs in humanized mice. *Proceedings of the National Academy of Sciences*, 110(41):16538–16543, 2013.

[38] J. Huang, G. Ofek, L. Laub, M. K. Louder, N. A. Doria-Rose, N. S. Longo, H. Imamichi, R. T. Bailer, B. Chakrabarti, S. K. Sharma, et al. Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. *Nature*, 2012.

[39] M. Jallow, Y. Y. Teo, K. S. Small, K. A. Rockett, P. Deloukas, T. G. Clark, K. Kivinen, K. A. Bojang, D. J. Conway, M. Pinder, et al. Genome-wide and fine-resolution association analysis of malaria in West Africa. *Nature Genetics*, 41(6):657–665, 2009.

[40] B. L. Jilek, M. Zarr, M. E. Sampah, S. A. Rabi, C. K. Bullen, J. Lai, L. Shen, and R. F. Siliciano. A quantitative basis for antiretroviral therapy for HIV-1 infection. *Nature Medicine*, 18(3):446–451, 2012.

[41] B. E. Johnson, T. Mazor, C. Hong, M. Barnes, K. Aihara, C. Y. McLean, S. D. Fouse, S. Yamamoto, H. Ueda, K. Tatsuno, et al. Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. *Science*, 343(6167):189–193, 2014.

[42] V. Jonsson, N. Matni, and R. M. Murray. Reverse engineering combination therapies for evolutionary dynamics of disease: An $\mathcal{H}_\infty$ approach. In *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, pages 2060–2065. IEEE, 2013.

[43] V. Jonsson, N. Matni, and R. M. Murray. Synthesizing combination therapies for evolutionary dynamics of disease for nonlinear pharmacodynamics. In *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pages 2352–2358. IEEE, 2014.

[44] V. Jonsson, A. Rantzer, and R. M. Murray. A scalable formulation for engineering combination therapies for evolutionary dynamics of disease. In *American Control Conference, (ACC), 2014*, pages 2771–2778. IEEE, 2014.

[45] M. G. Joyce, M. Kanekiyo, L. Xu, C. Biertümpfel, J. C. Boyington, S. Moquin, W. Shi, X. Wu, Y. Yang, Z.-Y. Yang, et al. Outer domain of hiv-1 gp120: antigenic optimization, structural malleability, and crystal structure with antibody vrc-pg04. *Journal of virology*, 87(4):2294–2306, 2013.

[46] R. Kantor, R. Shafer, S. Follansbee, J. Taylor, D. Shilane, N. Hurley, Leo, D. Phuong, D. Katzenstein, and W. J. Fesselb. Evolution of resistance to drugs in HIV-1-infected patients failing antiretroviral therapy. *AIDS*, 18(11):1503–1511, July 2004.

[47] F. Klein, A. Halper-Stromberg, J. A. Horwitz, H. Gruell, J. F. Scheid, S. Bournazos, H. Mouquet, L. A. Spatz, R. Diskin, A. Abadir, et al. HIV therapy by a combination of broadly neutralizing antibodies in humanized mice. *Nature*, 492(7427):118–122, 2012.

[48] S. Kobayashi, T. J. Boggon, T. Dayaram, P. A. Janne, O. Kocher, M. Meyerson, B. E. Johnson, M. J. Eck, D. G. Tenen, and B. Halmos. EGFR mutation and resistance of non–small-cell lung cancer to gefitinib. *New England Journal of Medicine*, 352(8):786–792, 2005.

[49] R. Kong, M. K. Louder, K. Wagh, R. T. Bailer, K. Greene, H. Gao, J. D. Taft, A. Gazumyan, C. Liu, M. C. Nussenzweig, et al. Improving neutralization potency and breadth by combining broadly reactive hiv-1 antibodies targeting major neutralization epitopes. *Journal of Virology*, 89(5):2659–2671, 2015.

[50] B. A. Larder and S. D. Kemp. Multiple mutations in HIV-1 reverse transcriptase confer high-level resistance to zidovudine (azt). *Science*, 246(4934):1155–1158, 1989.

[51] B. Lee, M. Sharron, L. J. Montaner, D. Weissman, and R. W. Doms. Quantification of CD4, CCR5, and CXCR4 levels on lymphocyte subsets, dendritic cells, and differentially conditioned monocyte-derived macrophages. *Proceedings of the National Academy of Sciences*, 96(9):5215–5220, 1999.

[52] M. Li, F. Gao, J. R. Mascola, L. Stamatatos, V. R. Polonis, M. Koutsoukos, G. Voss, P. Goepfert, P. Gilbert, K. M. Greene, et al. Human immunodeficiency virus type 1 Env clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies. *Journal of Virology*, 79(16):10108–10125, 2005.

[53] F. Lin, M. Fardad, and M. Jovanovic. Sparse feedback synthesis via the alternating direction method of multipliers. In *American Control Conference (ACC), 2012*, pages 4765–4770, 2012.

[54] J. Liu, A. Bartesaghi, M. J. Borgnia, G. Sapiro, and S. Subramaniam. Molecular architecture of native HIV-1 gp120 trimers. *Nature*, 455(7209):109–113, 2008.

[55] L. A. Loeb. A mutator phenotype in cancer. *Cancer research*, 61(8):3230–3239, 2001.

[56] R. Luo, M. J. Piovoso, J. Martinez-Picado, and R. Zurakowski. Optimal antiviral switching to minimize resistance risk in HIV therapy. *PloS One*, 6(11):e27047, 2011.

[57] R. M. Lynch, P. Wong, L. Tran, S. O'Dell, M. C. Nason, Y. Li, X. Wu, and J. R. Mascola. HIV-1 fitness cost associated with escape from the VRC01 class of CD4 binding site neutralizing antibodies. *Journal of Virology*, 89(8):4201–4213, 2015.

[58] L. Mansky and H. Temin. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *Journal of Virology*, 69(8):5087–94, 1995.

[59] N. Matni. Communication delay co-design in $\mathcal{H}_2$ decentralized control using atomic norm minimization. In *IEEE Conf. on Decision and Control, 2013*, 2013.

[60] N. Matni and V. Chandrasekaran. Regularization for design. In *IEEE 53rd Annual Conference on Decision and Control (CDC)*, pages 1111–1118. IEEE, 2014.

[61] T. McQuade, A. Tomasselli, L. Liu, V. Karacostas, B. Moss, T. Sawyer, R. Heinrikson, and W. Tarpley. A synthetic hiv-1 protease inhibitor with antiviral activity arrests hiv-like particle maturation. *Science*, 247(4941):454–456, 1990.

[62] L. M. Merlo, J. W. Pepper, B. J. Reid, and C. C. Maley. Cancer as an evolutionary and ecological process. *Nature Reviews Cancer*, 6(12):924–935, 2006.

[63] F. Michor, Y. Iwasa, and M. A. Nowak. Dynamics of cancer progression. *Nature Reviews Cancer*, 4(3):197–205, 2004.

[64] D. Montefiori. Meeting on development of monoclonal antibodies for HIV treatment and cure, unpublished.

[65] M. Montero, N. E. van Houten, X. Wang, and J. K. Scott. The membrane-proximal external region of the human immunodeficiency virus type 1 envelope: dominant site of antibody neutralization and target for vaccine design. *Microbiology and molecular biology reviews*, 72(1):54–84, 2008.

[66] H. Mouquet, L. Scharf, Z. Euler, Y. Liu, C. Eden, J. F. Scheid, A. Halper-Stromberg, P. N. Gnanapragasam, D. I. Spencer, M. S. Seaman, et al. Complex-type n-glycan recognition by potent broadly neutralizing HIV antibodies. *Proceedings of the National Academy of Sciences*, 109(47):E3268–E3277, 2012.

[67] D. G. Myszka, R. W. Sweet, P. Hensley, M. Brigham-Burke, P. D. Kwong, W. A. Hendrickson, R. Wyatt, J. Sodroski, and M. L. Doyle. Energetics of the HIV gp120-CD4 binding reaction. *Proceedings of the National Academy of Sciences*, 97(16):9026–9031, 2000.

[68] M. W. Nachman and S. L. Crowell. Estimate of the mutation rate per nucleotide in humans. *Genetics*, 156(1):297–304, 2000.

[69] F. Nimmerjahn and J. V. Ravetch. Antibody-mediated modulation of immune responses. *Immunological reviews*, 236(1):265–275, 2010.

[70] M. Nowak and R. M. May. *Virus dynamics: mathematical principles of Immunology and virology: mathematical principles of Immunology and virology*. Oxford university press, 2000.

[71] M. A. Nowak, R. M. May, and R. M. Anderson. The evolutionary dynamics of hiv-1 quasispecies and the development of immunodeficiency disease. *Aids*, 4(11):1095–1104, 1990.

[72] P. C. Nowell. The clonal evolution of tumor cell populations. *Science*, 194(4260):23–28, 1976.

[73] G. Ofek, B. Zirkle, Y. Yang, Z. Zhu, K. McKee, B. Zhang, G.-Y. Chuang, I. S. Georgiev, S. O'Dell, N. Doria-Rose, et al. Structural basis for HIV-1 neutralization by 2F5-like antibodies m66 and m66. 6. *Journal of Virology*, 88(5):2426–2441, 2014.

[74] A. Palsson and G. Gibson. Association between nucleotide variation in Egfr and wing shape in Drosophila melanogaster. *Genetics*, 167(3):1187–1198, 2004.

[75] M. Pancera, S. Shahzad-ul Hussan, N. A. Doria-Rose, J. S. McLellan, R. T. Bailer, K. Dai, S. Loesgen, M. K. Louder, R. P. Staupe, Y. Yang, et al. Structural basis for diverse n-glycan recognition by hiv-1–neutralizing v1–v2–directed antibody pg16. *Nature structural & molecular biology*, 20(7):804–813, 2013.

[76] W. Pao, V. A. Miller, K. A. Politi, G. J. Riely, R. Somwar, M. F. Zakowski, M. G. Kris, and H. Varmus. Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS Medicine*, 2(3):e73, 2005.

[77] A. S. Perelson, A. U. Neumann, M. Markowitz, J. M. Leonard, and D. D. Ho. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science*, 271(5255):1582–1586, 1996.

[78] R. Phillips, J. Kondev, J. Theriot, N. Orme, and H. Garcia. *Physical biology of the cell*. Garland Science New York, 2009.

[79] R. R. Picard and R. D. Cook. Cross-validation of regression models. *Journal of the American Statistical Association*, 79(387):575–583, 1984.

[80] L. H. Pinto, L. J. Holsinger, and R. A. Lamb. Influenza virus m 2 protein has ion channel activity. *Cell*, 69(3):517–528, 1992.

[81] E. J. Platt, K. Wehrly, S. E. Kuhmann, B. Chesebro, and D. Kabat. Effects of CCR5 and CD4 cell surface concentrations on infections by macrophagetropic isolates of human immunodeficiency virus type 1. *Journal of virology*, 72(4):2855–2864, 1998.

[82] A. Rantzer. Distributed control of positive systems. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 6608–6611, Dec. 2011.

[83] A. Rantzer. Distributed control of positive systems. In *IEEE Transactions on Automatic Control*. In preparation, 2014.

[84] A. Rantzer and B. Bernhardsson. Control of convex monotone systems. In *Decision and Control (CDC), 2014 IEEE 53nd Annual Conference on*. IEEE, 2014.

[85] B. Recht, M. Fazel, and P. A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.*, 52(3):471–501, Aug. 2010.

[86] R. R. Regoes, C. Wiuff, R. M. Zappala, K. N. Garner, F. Baquero, and B. R. Levin. Pharmacodynamic functions: a multiparameter approach to the design of antibiotic treatment regimens. *Antimicrobial agents and chemotherapy*, 48(10):3670–3676, 2004.

[87] D. I. S. Rosenbloom, A. L. Hill, S. A. Rabi, R. F. Siliciano, and M. A. Nowak. Antiretroviral dynamics determines HIV evolution and predicts therapy outcome. *Nature Biotechnology*, 18:1378–1385, June 2012.

[88] J. S Gach, D. P Leaman, and M. B Zwick. Targeting HIV-1 gp41 in close proximity to the membrane using antibody and other molecules. *Current topics in medicinal chemistry*, 11(24):2997–3021, 2011.

[89] R. Sadjadpour, O. K. Donau, M. Shingai, A. Buckler-White, S. Kao, K. Strebel, Y. Nishimura, and M. A. Martin. Emergence of gp120 V3 variants confers neu-

tralization resistance in an R5 simian-human immunodeficiency virus-infected macaque elite neutralizer that targets the N332 glycan of the human immunodeficiency virus type 1 envelope glycoprotein. *Journal of Virology*, 87(15):8798–8804, 2013.

[90] M. E. S. Sampah, L. Shen, B. L. Jilek, and R. F. Siliciano. Dose-response curve slope is a missing dimension in the analysis of HIV-1 drug resistance. *Proceedings of the National Academy of Sciences*, 108(18):7613–7618, 2011.

[91] L. Scharf, J. F. Scheid, J. H. Lee, A. P. West, C. Chen, H. Gao, P. N. Gnanapragasam, R. Mares, M. S. Seaman, A. B. Ward, et al. Antibody 8ANC195 reveals a site of broad vulnerability on the HIV-1 envelope spike. *Cell Reports*, 7(3):785–795, 2014.

[92] J. F. Scheid, H. Mouquet, B. Ueberheide, R. Diskin, F. Klein, T. Y. Oliveira, J. Pietzsch, D. Fenyo, A. Abadir, K. Velinzon, et al. Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science*, 333(6049):1633–1637, 2011.

[93] J. Schymkowitz, J. Borg, F. Stricher, R. Nys, F. Rousseau, and L. Serrano. The FoldX web server: an online force field. *Nucleic acids research*, 33(suppl 2):W382–W388, 2005.

[94] L. V. Sequist, J.-C. Soria, J. W. Goldman, H. A. Wakelee, S. M. Gadgeel, A. Varga, V. Papadimitrakopoulou, B. J. Solomon, G. R. Oxnard, R. Dziadziuszko, et al. Rociletinib in egfr-mutated non–small-cell lung cancer. *New England Journal of Medicine*, 372(18):1700–1709, 2015.

[95] M. Shingai, Y. Nishimura, F. Klein, H. Mouquet, O. K. Donau, R. Plishka, A. Buckler-White, M. Seaman, M. Piatak, J. D. Lifson, et al. Antibody-mediated immunotherapy of macaques chronically infected with SHIV suppresses viraemia. *Nature*, 503(7475):277–280, 2013.

[96] T. Tanaka and C. Langbort. The bounded real lemma for internally positive systems and H-infinity structured static state feedback. *IEEE Transactions on Automatic Control*, 56(9):2218–2223, 2011.

[97] M. D. Thakur, F. Salangsang, A. S. Landman, W. R. Sellers, N. K. Pryer, M. P. Levesque, R. Dummer, M. McMahon, and D. D. Stuart. Modelling vemurafenib resistance in melanoma reveals a strategy to forestall drug resistance. *Nature*, 494(7436):251–255, 2013.

[98] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.

[99] S. Verstovsek, R. A. Mesa, J. Gotlib, R. S. Levy, V. Gupta, J. F. DiPersio, J. V. Catalano, M. Deininger, C. Miller, R. T. Silver, et al. A double-blind, placebo-controlled trial of ruxolitinib for myelofibrosis. *New England Journal of Medicine*, 366(9):799–807, 2012.

[100] X. Wei, J. M. Decker, S. Wang, H. Hui, and e. a. Kappes, John C. Antibody neutralization and escape by HIV-1, volume = 422, year = 2003. *Nature*, pages 307–312.

[101] A. P. West, L. Scharf, J. Horwitz, F. Klein, M. C. Nussenzweig, and P. J. Bjorkman. Computational analysis of anti–HIV-1 antibody neutralization panel data to identify potential functional epitope residues. *Proceedings of the National Academy of Sciences*, 110(26):10598–10603, 2013.

[102] K. B. Wood, K. C. Wood, S. Nishida, and P. Cluzel. Uncovering scaling laws to infer multidrug response of resistant microbes and cancer cells. *Cell Reports*, 6(6):1073–1084, 2014.

[103] S. Wright. *The roles of mutation, inbreeding, crossbreeding, and selection in evolution*, volume 1. na, 1932.

[104] X. Wu, T. Zhou, J. Zhu, B. Zhang, I. Georgiev, C. Wang, X. Chen, N. S. Longo, M. Louder, K. McKee, et al. Focused evolution of HIV-1 neutralizing antibodies

revealed by structures and deep sequencing. *Science*, 333(6049):1593–1602, 2011.

[105] B. Zhao, M. T. Hemann, and D. A. Lauffenburger. Intratumor heterogeneity alters most effective drugs in designed combinations. *Proceedings of the National Academy of Sciences*, 111(29):10773–10778, 2014.

[106] B. Zhao, J. R. Pritchard, D. A. Lauffenburger, and M. T. Hemann. Addressing genetic tumor heterogeneity through computationally predictive combination therapy. *Cancer discovery*, 4(2):166–174, 2014.

[107] T. Zhou, J. Zhu, X. Wu, S. Moquin, B. Zhang, P. Acharya, I. S. Georgiev, H. R. Altae-Tran, G.-Y. Chuang, M. G. Joyce, et al. Multidonor analysis reveals structural elements, genetic determinants, and maturation pathway for hiv-1 neutralization by vrc01-class antibodies. *Immunity*, 39(2):245–258, 2013.

[108] R. Zurakowski. Nonlinear observer output-feedback MPC treatment scheduling for HIV. *Biomedical Engineering Online*, 10(1):40, 2011.

[109] M. B. Zwick, R. Jensen, S. Church, M. Wang, G. Stiegler, R. Kunert, H. Katinger, and D. R. Burton. Anti-human immunodeficiency virus type 1 (HIV-1) antibodies 2F5 and 4E10 require surprisingly few crucial residues in the membrane-proximal external region of glycoprotein gp41 to neutralize HIV-1. *Journal of Virology*, 79(2):1252–1261, 2005.