Identification of Cancer Cell Population Dynamics Leveraging the Effect of Pre-Treatment for Drug Schedule Design

Marius Wiggert[1], Megan Turnidge[2], Zoe Cohen[1], Ellen M. Langer[2], Rosalie C. Sears[2], Margaret P. Chapman[3] and Claire J. Tomlin[1]

***Abstract*— Sequences of different drugs have shown potential to improve treatment strategies for cancer. Typical switched system approaches model the population dynamics of each drug independently, not rigorously considering the effects of pre-treatment or drug-drug interactions. In this paper, a general model family incorporating pre-treatment effects and biological domain knowledge is proposed, and a model from this family is identified by using a novel experimental data set of two-drug sequences. Leveraging the data, a simulator for the cell population dynamics under sequences of up to nine drugs is developed and used to empirically evaluate the performance of a set of closed-loop drug scheduling controllers. We used the controllers to identifying promising drug schedules *in silico* and evaluated them *in vitro*. The experiments validated the effectiveness of the identified schedules in reducing the number of living cells to less than 10% of the initial. While only treating with certain toxic drugs achieves similar effectiveness, the schedules use toxic drugs for significantly shorter times which likely reduces toxicity to non-cancer cells.**

I. Introduction

Understanding and predicting the response of cancer cell populations or other sources of disease, such as bacteria or viruses, to different treatments is important for improving therapeutic strategies. Therefore, the mathematical modeling of cancer cell populations and their response to varying doses of targeted drugs is an active research area [1]. It is known that combinations of drugs acting synergistically can improve the effectiveness and specificity of therapies [2], [3]. Moreover, past studies report that these synergistic effects depend on the time and sequence in which the drugs are applied [4], [5]. Hence, the application of multiple drugs in a *sequence* has the potential to effectively stabilize the cancer population while minimizing toxicity [6] and mitigating the risk of developing drug resistance [5], [7].

A core challenge for realizing this potential is to identify promising schedules amongst the combinatorial amount of drug combinations and sequence timings. Prior experiments focus on identifying synergistic or antagonistic effects of drugs with two-drug sequences by comparing the final cell counts. The drug pairs are selected based on biological domain knowledge, or recently by leveraging high throughput screening of combinations of 100 drugs [8]. However, even with high throughput screening, this trial-and-error approach is limited in evaluating the combinatorial space of possible drug schedules. Treating the cell population as a dynamical system enables the use of control tools to identify promising schedules *in silico* which can then be tested *in vitro*.

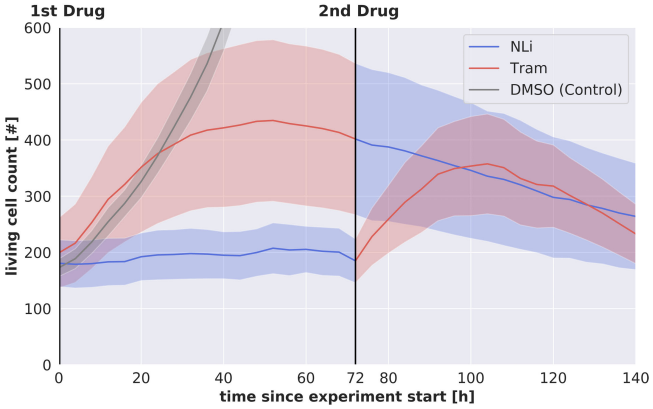

Fig. 1: Time-series data from experiments of the response of breast cancer cell populations to two drugs (Trametinib and NLi) applied sequentially (t=0, t=72h) compared to exponentially growing untreated cells (DMSO). The lines represent the mean of 6 replicate wells, and the shaded areas the standard deviation. Observe, that the population dynamics after NLi treatment are different when applied at time 0 compared to after 72 hours of Trametinib pre-treatment.

Our previous work modeled the cancer cell population response to a sequence of drugs as a switched dynamical system, using data of the response to a single drug to identify the dynamics [9]. One of the key assumptions for the resulting drug schedules to asymptotically reduce the cell population to zero was that the error due to neglecting drug-drug interactions is sufficiently small. In this paper, we model the indirect interaction of drugs through pre-treating cells with one drug, before treating with a second and thereby improve upon our former approach by using newly available time-series data of a population's response to sets of two drugs applied sequentially. As expected, the data indicates that the population's response to a given drug can be altered by pre-treatment with a different drug (Fig. 1).

Developing mathematical models and control strategies of the response of cancer cell populations to drugs in sequence, taking into account the effect of pre-treatment, is an open research area. In this paper, we use data of breast cancer cell populations treated with 9 drugs in various two-drug

[1]M.W., Z.C., and C.J.T. are with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, USA. For inquiries contact: mariuswiggert@berkeley.edu

[2]M.T., E.M.L., and R.C.S. are with the Department of Molecular and Medical Genetics, Oregon Health and Science University, USA.

[3] M.P.C. is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, Canada.

The authors gratefully acknowledge the support of the NCI CSBC program and the German Academic Exchange Service (DAAD).

treatment schedules, where each condition lasts 72 hours. We use this novel data set to quantify how the cell population dynamics differ when drugs are applied in different orders, develop mathematical models that incorporate these differences, and leverage these models for simulation and the identification of promising drug schedules.

Related Work. Previous work has been done on modeling cell populations as linear or nonlinear systems [6], [7]. Also, many strategies have been explored to mathematically identify optimal treatment regimens to manage cancer, HIV, and other illnesses [10], [11] where it is especially important to minimize drug dose. Control strategies that result in more efficient treatment (for example, less total drug volume applied) for a single drug can be found by solving linear matrix inequalities using convex optimization techniques [10]. The type of model used for cell populations or dynamics can also vary widely, including biological networks [10], [12], nonlinear models [11], [13], systems of linear ordinary differential equations [14], or stochastic hybrid models [15]. Previous work has shown that combinations of drugs can be particularly effective for avoiding the onset of drug resistance in cancer treatment by exploiting phenotypic state transitions [5]. Formulating treatment as a control problem has the potential to mitigate the trade-off of drug efficacy versus toxicity. The problem of modeling the effect of applying a set of drugs in a sequence is well suited for representation as a switched system. There has been extensive work on the stability and control of these systems [16], [17].

Contributions. Motivated by the effect of pre-treatment, we suggest a general family of dynamics models (Sec. III) that incorporates both pre-treatment effects and various types of domain knowledge. The most suitable model of this family is identified (Sec. IV) using novel experimental data of two-drug schedules (Sec. II). We use our models of cell population dynamics under nine drugs to build a simulator on which drug schedules can be tested *in silico* (Sec. V). Further, we present initial work on designing closed-loop drug schedule controllers (Sec. V) which we use to identify promising drug schedules *in silico*. We evaluate these schedules *in vitro* (in cell cultures in a laboratory). The methods we develop can be applied to any switched system in which the previous mode influences the dynamics in the consecutive mode, if sufficient data of sequences is available.

II. Cancer Cell Experiments and Data Set

In this section, we describe the cancer cell experiments and the resulting time-series data set that is used for model identification, which will be presented in (Sec. IV).

Sequential drug treatments were applied to the breast cancer cell line SUM149PT, with the first drug administered at time zero and the second drug at time 72 hours, after a washout of the medium containing the first drug (see Fig. 1). To measure the effect of treatment on cell population growth, wavelength-specific images were taken every 4 hours using the IncuCyte ZOOM imaging and digital segmentation system (Essen BioScience, Ann Arbor, USA). The two available measurements are: 1) the number of living cells, as the cells used were infected with a lentivirus labeling each nucleus with red fluorescence and 2) the number of dying cells using a green fluorescent marker for detecting Caspase 3/7 cleavage during apoptosis i.e. cell death (only for schedules involving the drugs BEZ235 and JQ-1).

46 unique two-drug treatment schedules were conducted experimentally, each with 6-10 physically separate replicate *wells* of cell populations, which we assume to be independent and identically distributed. Using the drug vehicle dimethyl sulfoxide (DMSO) as the baseline condition which simulates normal untreated cell growth, the following nine treatment schedules were applied for each pair of drugs: drug 1/DMSO, drug 1/drug 1, drug 1/drug 2, drug 1/combination, drug 2/DMSO, drug 2/drug 2, drug 2/drug 1, drug 2/combination, and combination/combination, where "combination" denotes both drug 1 and drug 2. Our experiments include the following pairs of drugs: 1) BEZ235 (a PI3K/mTOR inhibitor) & JQ-1 (a BET inhibitor), 2) Trametinib (a MEK inhibitor) & a nuclear lamin inhibitor (NLi), 3) Trametinib & a PARP inhibitor, 4) Trametinib & BRD4 inhibitor A, 5) Trametinib & BRD4 inhibitor B, 6) Trametinib & BRD4 inhibitor C.

The time-series data set was constructed by exporting the measured number of cells from each well at a given time point when an image was taken. This data set consists of 46 treatment schedules with 6-10 replicate wells and 36 time-series data points per well.

III. Mathematical Modeling

In this section, we present a family of mathematical models and the rationale for the modeling choices.

A. Model Family

We represent the evolution of a cancer cell population under a drug schedule as a *an uncertain, discrete-time, time-invariant, switched linear auto-regressive* dynamical system, with the state at time step $t \in \mathbb{N}_+$ represented by the n dimensional non-negative state vector $x_t \in \mathbb{R}^n_+$. In this paper we consider $n \in \{1, 2\}$ and the discrete time steps are 4h apart (i.e. $t_{real}{=}t \cdot 4h$). For $n{=}1$ the state vector x_t contains the numbers of living cells $x_t{=}$[#living], and for $n{=}2$ additionally the number of dying cells $x_t{=}[\#\text{living}, \#\text{dying}]^T$, taken from the available measurements (Sec. II). We chose this high-level state representation as it is useful for control, explicitly abstracting away from the complexity of modelling cell-level Pharmacodynamics. The current state x_t depends on the p previous states $(x_{t-1}, x_{t-2}, \ldots, x_{t-p})$. A *drug schedule*, denoted by σ, is a mapping from time step t to the most recent drug applied at or before time step t, hence $\forall t \ \sigma[t] \in D, t \in \mathbb{N}_+$, where D is the set of available drugs. The model also accounts for additive process noise, a multiplicative term that represents the uncertain effect of pre-treatment on the cell population dynamics, and temporal differences in drug activity. An element of the family of mathematical models over a finite time horizon of T time steps is denoted by a tuple (N_{tw}, p, n, C) with the number of time windows per drug N_{tw}, the number of states n, the

auto-regressive order p, and the constraints C. The models are of the following form:

$$x_t = \sum_{i=1}^{p} A^i_{\sigma[t]} \cdot \xi^i_{\sigma[t],t} \cdot x_{t-i} + \eta_{\sigma[t],t}, \tag{1}$$

where the system matrices $A^i_{\sigma[t]} \in \mathbb{R}^{n\times n}$ represents the influence of the i^{th} previous state on the current state after application of drug $\sigma[t]$ in a *treatment-naive* setting. *Treatment-naive* is defined as the condition in which the cancer cell population has not been exposed to treatment before the current treatment, namely $\sigma[t]$ is the first drug applied. In contrast, *post-treatment* refers to the condition in which the cancer cell population has been treated with another drug before the current drug. For $n = 2$, $A^i_{\sigma[t]}$ is a matrix of the form

$$A^i_{\sigma[t]} := \begin{bmatrix} A^i_{\sigma[t],00} & A^i_{\sigma[t],01} \\ A^i_{\sigma[t],10} & A^i_{\sigma[t],11} \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

where $A^i_{\sigma[t],kl}$ for $k, l \in \{0, 1\}$ models the influence of the i^{th} previous number of living ($l = 0$)/dying ($l = 1$) cells on the number of current living ($k = 0$)/dying ($k = 1$) cells. The drug-specific process noise is represented by $\eta_{\sigma[t],t} \in \mathbb{R}^n$. The multiplicative term $\xi^i_{\sigma[t],t} \in \mathbb{R}$ represents the uncertainty in the change of the system dynamics under a drug $\sigma[t]$ from the treatment-naive to post-treatment conditions. Hence, we refer to $\xi^i_{\sigma[t],t}$ as the *post-treatment dynamics uncertainty*. Note that the term $\xi^i_{\sigma[t],t}$ is needed because our data does not contain all possible drug sequences. We introduce constraints C and time windows N_{tw} as additional parameters that allow us to incorporate domain knowledge.

Constraints C on the system matrices $(A^i_{\sigma[t]})^p_{i=1}$ encode how states can affect each other. Observe that for $x_t \in \mathbb{R}$ the matrices $(A^i_{\sigma[t]})^p_{i=1}$ are scalars hence we do not use constraints in that case ($C = \{\text{None}\}$). For $n = 2$, $x_t \in \mathbb{R}^2$ we consider three different sets of constraints on the system matrices: a) no constraints, b) $A^i_{\sigma[t],01} = 0$, meaning the number of dying cells does not influence the future number of living cells, and c) $A^i_{\sigma[t],01} = 0$ and $A^i_{\sigma[t],11} \geq 0$, where the latter inequality forces the quantity of dying cells to positively influence the quantity of future dying cells.

Time windows allow us to model the different modes of drug activity. We divide the available 72 hours of respective naive- and post-treatment measurements after drug application into up to three equally-spaced time windows $N_{\text{tw}} \in \{1, 2, 3\}$. $N_{\text{tw}} = 1$ represents one mode of drug activity: $T_{\text{w}} = [0, 18] = \{0, 1, \ldots, 18\}$ for the measurements from 0h to 72h, for $N_{\text{tw}} = 2$ we have: $T^1_{\text{w}} = [0, 9]$ and $T^2_{\text{w}} = [9, 18]$ and lastly for $N_{\text{tw}} = 3$: $T^i_{\text{w}} = [6(i-1), 6i]\ \forall i \in \{1, 2, 3\}$.

Assumptions: We make the following modeling assumptions regarding η and ξ. The drug-specific process noise $\eta_{\sigma[t],t}$ is assumed to normally distribution with zero mean ($\mu = 0_n$) and a drug-specific diagonal covariance matrix $\Lambda_{\sigma[t]} \in \mathbb{R}^{n\times n}$. After applying the first drug there is no post-treatment dynamics uncertainty as the system matrices $A^i_{\sigma[t]}$ represent treatment-naive dynamics, consequently $\xi^i_{\sigma[t],t} = 1$ for $i = 1, 2, \ldots, p$ for all t following the first drug treatment prior to the second drug treatment. From the first *drug switch*, at the time step of the second drug treatment, onward $\xi_{\sigma[t],t} := (\xi^1_{\sigma[t],t}, ..., \xi^p_{\sigma[t],t})$ takes on some value drawn at the respective switching time step t_s from a bounded drug-specific distribution, denoted by $F_{\sigma[t]}$, to reflect the uncertain change in the drugs treatment-naive dynamics due to pre-treatment. A *drug switch* is defined as $\sigma[t_s - 1] \neq \sigma[t_s]$. Hence, $A^i_{\sigma[t_s]} \cdot \xi^i_{\sigma[t_s],t_s}$ represents the post-treatment dynamics of the cancer cell population. $\xi_{\sigma[t],t}$ is drawn only at the respective switching time t_s. Consequently, $\xi_{\sigma[t_s],t_s} = \xi_{\sigma[j],j}$ for all j after t_s but prior to the next drug switch. We assume the effect of pre-treatment to be modeled completely by the multiplicative term $\xi_{\sigma[t],t}$ (rational Sec. III-B).

We estimate the treatment-naive and post-treatment matrices $(A^i_{\sigma[t]})^p_{i=1}$, the covariance matrix $\Lambda_{\sigma[t]}$, and the distribution of the post-treatment dynamics uncertainty $F_{\sigma[t]}$ for each drug $\sigma[t] \in D$ using our time-series data set.

Note that because our data consists of two-drug schedules to fit $F_{\sigma[t]}$, which only models the interaction between the current and the previous drug. This implicitly assumes that in a sequence of more than two drugs e.g. A-B-C, the dynamics under C are only influenced by the previous drug B and not by A. This is biologically justified for some drug mechanisms (e.g. the drug effect vanishes over time) but likely does not hold for all mechanisms. However, it is the best we can do with the data available.

B. Modeling Rationale

The choice of the model family is justified as follows:

(1) Cell divisions are inherently discrete and without nutrient limits population growth is exponential. Linear discrete-time auto-regressive models are sufficiently general to represent such higher-order dynamical behavior while remaining simple enough for parameter estimation, practical computation, and efficient control.

(2) We decouple $\eta_{\sigma[t],t}$ and $\xi_{\sigma[t],t}$ since these terms represent two different sources of uncertainty: standard additive process noise and post-treatment dynamics uncertainty, respectively. The post-treatment effect is uncertain as data is only available for certain drug-drug sequences, hence to predict likely effects for arbitrary schedules we model this uncertainty as a distribution (details in Sec. IV-B.2). We chose to model the dynamics uncertainty as multiplicative term $\xi_{\sigma[t],t}$ instead of more expressive alternatives such additive terms for each matrix element, because the multiplicative method requires only one parameter, and our data set is not currently large enough to identify 4 unique parameters required for the additive method.

(3) The model family allows for the incorporation of domain knowledge. Knowledge about how the past number of living/dying cells influence the future numbers can be captured in the form of constraints on elements of the system matrices $A^i_{\sigma[t]}$. Moreover, pharmacological knowledge about the temporal variation of drug activity can be incorporated by having multiple, temporally sequenced models of type (1) that represent sequential modes of drug activity. We consider two effects, firstly drugs often have a temporal delay until maximum drug activity is observed and secondly that drug

activity decreases after some time [18]. Hence we divide the 72 hour time frame over which a specific drug is active in $N_{\text{tw}} \in \{1, 2, 3\}$ *time windows*.

IV. System Identification

This section details the identification procedure and respective results for the system matrices $(A^i_\delta)^p_{i=1}$, the process noise covariance matrix Λ_δ, and the distribution of the post-treatment dynamics uncertainty F_δ for each drug $\delta \in D$.

A. System Matrices

1) System Identification Methodology: We identify the system matrices $(A^i_\delta)^p_{i=1}$ for a specific drug $\delta \in D$ on a set of training wells W_{Train} using a least-squares loss. Note that for every drug δ there is one treatment-naive condition and multiple post-treatment conditions. We denote their respective system matrices by $A^i_{\delta,\text{Naive}}$ and $A^i_{\delta,\text{Post}}$.

The respective drug- and condition-specific experimental measurements are split randomly into the two disjoint sets, W_{Train} and W_{Test}, where the latter contains two wells and the former contains 4-8 wells ($\approx$ 66-80% of all available wells $W_{\text{Total}} = W_{\text{Train}} \cup W_{\text{Test}}$). For example, we use the experimental data summarized in the upper curve of Fig. 1 until $T_s = 18$ (72 hours) to estimate $(A^i_{\delta,\text{Naive}})^p_{i=1}$ for $\delta =$ Trametinib, whereas we use the data in the lower curve after $t_s = 18$ (72 hours) to estimate $(A^i_{\delta,\text{Post}})^p_{i=1}$ for the NLi post-treatment condition. Under our assumption of additive zero-mean Gaussian process noise, the least-squares loss can be derived from the maximum likelihood inference principle [19]. Recall that an element of the model family is fully specified by a tuple of four variables (N_{tw}, p, n, C). For a given drug $\delta \in D$ with data from the set W_{Train}, we fit the model for the time window $T_{\text{w}} = [T_{\text{L}}, T_{\text{U}}]$ by solving the following least-squares regression problem:

$$(A^i_\delta)^p_{i=1} = \underset{(A^i_\delta)^p_{i=1}}{\arg\min} \sum_{w \in W_{\text{Train}}} \sum_{t=T_{\text{L}}+p}^{T_{\text{U}}} \|x_{t,w} - \sum_{i=1}^{p} A^i_\delta \cdot x_{t-i,w}\|_2^2,$$

where $(A^i_\delta)^p_{i=1}$ are subject to the set of constraints C, $x_{t,w} \in \mathbb{R}^n_+$ is the measurement at time step t from well w, and T_{L} and T_{U} are the lower and upper bounds for the given time window, respectively. Observe that $\xi_{\delta,t}$ is not present above. The estimation of $\xi_{\delta,t}$ will be presented in Sec. IV-B.

After we fitted the drug-specific naive and post-treatment system matrices to the training data W_{Train}, we evaluated how well the resulting model generalizes to unseen test data W_{Test} by calculating the mean-squared-error (MSE) on the numbers of living cells in the respective time window. The numbers of dying cells were not always measured, and quantifying the numbers of living cells is more important for control, as we wish to drive these cells to zero. We calculate the MSE between the measured counts and k-step predictions using the identified system matrices. To compute k-step predictions, p values of the time-series are taken and the identified dynamics (1) are used to predict the k-following counts of living cells. Fig. 2 shows the predictions of the (N_{tw}=1, p=2, n=1, C=None) model for $k = 3$. We want to choose a high k because the simulator and controllers,

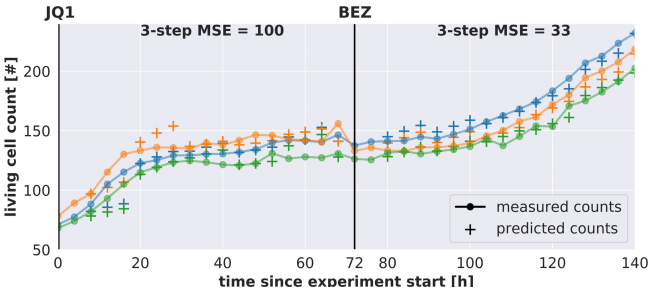

Fig. 2: The measured number of living cells from three wells of the JQ1/BEZ experiment data set (JQ1 given at t=0, BEZ at t=18 i.e. 72h). Additionally, the 3-step predictions of the second-order auto-regressive model with one time window per drug (N_{tw}=1, p=2, n=1, C=None) that has been fitted on the other seven wells are shown.

which will be presented in Sec. V, rely on predicting the cell population dynamics over multiple time steps. We choose $k = 3$ as our metric of comparison as three steps is the maximum amount to allow for a fair comparison (i.e. with $N_{\text{tw}} = 3, p = 3$ the time window is six steps long).

To get the best possible estimate of generalization performance, cross-validation for all combinations of two test wells out of W_{Total} is performed.

2) System Identification Results: This section presents the cross-validation results on the experimental data set and their interpretation to justify the selection of a suitable model from the family of system models (Sec. III-A). A total of 36 models, each specified by a tuple (N_{tw}, p, n, C), were evaluated by varying the following four parameters: the number of time windows per drug $N_{\text{tw}} \in \{1, 2, 3\}$, the number of states $n \in \{1, 2\}$, the auto-regressive order $p \in \{1, 2, 3\}$, and the constraints $C \in \{\text{None}, A^i_{01} = 0, A^i_{01} = 0 \,\&\, A^i_{11} \geq 0\}$.

The cross-validation results presented are aggregated across a subset of two-drug schedule experiments for which the $x_t \in \mathbb{R}^2$, because the number of dying cell measurement is not available for all experiments (Sec. II). We fit a (N_{tw}, p, n, C)-model to both the treatment-naive part of the experiment (i.e., from 0 hours to 72 hours), and to the post-treatment part (i.e., from 72 hours to 144 hours). For example, Fig. 2 shows the performance of the ($N_{\text{tw}} = 1$, $p = 2$, $n = 1$, $C =$ None)-model on three *unseen* test wells for the JQ1-then-BEZ experiment.

The cross-validation results are summarized in Fig. 3. While they are over a subset of experiments, the same analysis was also performed with n=1 over all 46 experiments. Moreover, to study the best model for individual two-drug schedule experiments (e.g. for JQ1/BEZ a specific model tuple might perform well but not for the BEZ/JQ1 experiment) the analysis was also performed for the experiments individually. While the exact 3-step MSE values are different, the qualitative conclusions are consistent, and we summarize these conclusions in the following.

(1) Our experimental results indicate that increasing the number of time windows N_{tw} reduces the testing 3-step MSE by an order of magnitude. This result is aligned with the biological domain knowledge that drugs have multiple phases

Fig. 3: Goodness of fit in terms of 3-step prediction mean-squared-error (MSE) cross validation for different models (N_{tw}, p, n, C) on the JQ1, BEZ, DMSO combination data set. $N_{\text{tw}} \in \{1, 2, 3\}$ is the number of time-windows per drug, $n \in \{1, 2\}$ the number of states, $p \in \{1, 2, 3\}$ the auto-regressive order, and $C \in \{\text{None}, \text{constrained}\}$. The y-axis is in logarithmic scale.

of effectiveness (Sec. III-B).

(2) We found that increasing the auto-regressive order from $p = 1$ to $p = 2$ essentially halves the MSE. Increasing the order from $p = 2$ to $p = 3$ does not yield a clear benefit, since MSE decreases for $N_{\text{tw}} = 2$ but increases for $N_{\text{tw}} = \{1, 3\}$. The latter suggests overfitting of the model.

(3) Our results show that constraining the system matrices for $n = 2$ has a small effect on MSE. We observe that out of the three constraints studied $C \in \{\text{None}, A^i_{01} = 0, A^i_{01} = 0 \,\&\, A^i_{11} \geq 0\}$, the latter two result in the same MSE, hence the $A^i_{11} \geq 0$ constraint is not active and only $A^i_{01} = 0$ matters. The constraint supports generalization for $p \geq 1$, $N_{\text{tw}} \geq 1$ but hurts performance for the other models.

(4) We found that increasing the number of states $n \in \{1, 2\}$ seems to have small and largely deteriorating effects on MSE. Only for $p = 1$ generalization performances improves when modeling the number of dying cells ($n = 2$).

In summary, increasing the auto-regressive order p and the number of time windows N_{tw} improve the model performance, while the effects of n and C are small and ambiguous. An important aspect to consider when choosing a model is the balance between the goodness of fit (measured by MSE) and the complexity of the dynamics model (in terms of the number of parameters N_{param}), which is the classic bias-variance trade-off. The number of parameters for the system matrices is $N_{\text{param}} = N_{\text{tw}} \cdot p \cdot n^2 - n \cdot I_{\text{constrained}}$, where $I_{\text{constrained}} = 1$ if C contains $\{A^i_{01} = 0\}$ and zero otherwise. We choose $n = 1$ and thereby $C = \text{None}$ because $n = 2$ does not consistently improve the fit, and for many drugs only $n = 1$ measurements are available. We choose $p = 2$ because our results indicate that for $p = 3$ the models overfit to the training data as discussed above. For the number of time windows, although $N_{\text{tw}} \in \{2, 3\}$ has superior performance, we choose $N_{\text{tw}} = 1$ since this reduces the complexity of the model for simulation and control (Sec. V). We will investigate models capturing the time-dependence of drug activity in future work. Hence, we choose (N_{tw}=1, p=2, n=1, C=None), which will be considered fixed hereafter.

B. Estimation and Results for Λ_δ and F_δ

1) Process Noise Covariance Matrix Λ_δ: The n process noise variances v_j $j \in \{1, 2\}$ as the elements of the diagonal Covariance Matrix $\Lambda_\delta \in \mathbb{R}^{n \times n}$ for a specific drug $\delta \in D$ are estimated by computing the residuals

$$r_{w,t} := x_{t,w} - \sum_{i=1}^{p} A^i_\delta \cdot x_{t-i,w} \tag{2}$$

Using the system matrices $(A^i_\delta)^p_{i=1}$ identified with the procedure described in Sec. IV-A. The residuals $r_{w,t}$ for all $w \in W_{\text{Total}}$ and across the respective time windows on which the model was fitted $t \in T_{\text{w}} = [T_{\text{L}}, T_{\text{U}}]$ resemble a distribution, on which for each state $j \in \{1, 2\}$, a normal distribution is fitted to obtain the variances v_j.

2) Post-treatment Dynamics Uncertainty Distribution F_δ: The multiple sets of post-treatment dynamics system matrices $(A^i_{\delta,\text{Post}})^p_{i=1}$ of a drug can be estimated using the procedure described in Sec. IV-A. We can think of the matrices $(A^i_{\delta,\text{Post}})^p_{i=1}$ as multi-modal, each mode corresponding to a different pre-treatment drug. These different modes can be observed in Fig. 4 which shows the different identified dynamics comparing treatment-naive PARPi, PARPi-after-Trametinib, and PARPi-after-PARPi.

Our experimental data for a specific drug δ only provides measurements for certain post-treatment conditions (drug-drug sequences). Hence, to estimate the distribution of likely post-treatment dynamic shifts F_δ, we make two assumptions: a) the post-treatment dynamics of δ for other drugs or application times are similar to the ones observed in our experiments, and b) the post-treatment dynamics for unseen post-treatment conditions are unknown, so ξ_{δ,t_s} is sampled at each drug switch t_s to estimate the dynamics $(A^i_{\delta,\text{Post}})^p_{i=1} = (A^i_{\delta,\text{Naive}} \cdot \xi_\delta)^p_{i=1}$. These assumptions are necessary due to limited data and are reflected in the estimation procedure for F_δ, which is done via bootstrapping. The two steps for obtaining $N_{\delta,\text{post}}$ samples of ξ_δ, where $N_{\delta,\text{post}}$ is the number of experiments in which δ is given as second drug are:

1) Estimate $(A^i_{\delta,\text{Naive}})^p_{i=1}$ and the $N_{\delta,\text{post}}$ different $(A^i_{\delta,\text{Post}})^p_{i=1}$ using bootstraps on the wells of the treatment naive and post-treatment conditions experiments.
2) Calculate for each of the $N_{\delta,\text{post}}$ post-treatment matrices $(A^i_{\delta,\text{Post}})^p_{i=1}$ an $\xi_\delta = (\xi^1_\delta, \ldots, \xi^p_\delta)^T$ using $\xi^i_\delta = \arg\min_{\xi^i_\delta} \|A^i_{\delta,\text{Naive}} \cdot \xi^i_\delta - A^i_{\delta,\text{Post}}\|^2_2$ for $i \in \{1, \ldots, p\}$

Repeating this procedure 1000 times yields a distribution of $1000 \cdot N_{\delta,\text{post}}$ samples of the post-treatment dynamics uncertainty ξ_δ to which we fit a kernel density estimator to obtain F_δ. We sample from this kernel density estimator in the simulator, which will be presented in Sec. V.

3) Estimation Results of Λ_δ and F_δ: We identified the empirical covariance Λ_δ and distribution F_δ for each drug δ using our chosen model of ($N_{\text{tw}} = 1$, $p = 2$, $n = 1$, $C =$ None). This was done to validate the Gaussian assumptions on η_δ and get insight into the post-treatment uncertainty distribution $\xi_\delta \sim F_\delta$. We found that η_δ is Gaussian for all drugs mostly with $\mu \approx 0$ and varying variances. For

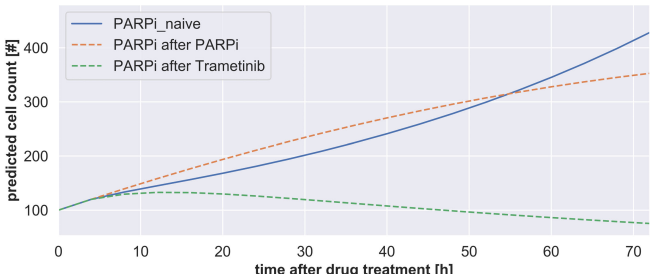

Fig. 4: The forward predictions of the fitted models of PARPi treatment-naive and two PARPi post-treatment conditions. Note significant variation in the population dynamics.

$\xi_{\delta,t} = (\xi^1_{\delta,t}, \xi^2_{\delta,t}) \sim F_\delta$ the observed empirical F_δ shows a strong correlation between ξ^1_δ and ξ^2_δ and is multi-modal for most drugs. This aligns with the observed difference in post-treatment dynamics dependent on the applied pre-treatment drug for which Fig. 4 shows an example.

V. Design of Drug Scheduling Controllers

This work aims is to identify promising treatment strategies *in silico*, thereby guiding laboratory research on future experiments and improving on the current trial-and-error paradigm to discover drug synergies. This section presents initial work in this direction. First, the design of our simulator to test drug schedules is detailed. Then a set of approaches for drug scheduling controllers are described and evaluated *in silico*. Lastly, the *in vitro* results of promising drug schedules identified with the controller *in silico* are presented.

A. Drug Schedule Simulator

The simulator models the number of living cells of a cancer cell population step-wise with $\Delta_t = 4h$, at every step $\sigma[t]$ the newly or last applied drug is given as input. Given a two measurement initial condition, the number of living cells evolves according to Eq. (1) with system matrices and drug-specific additive $\eta_{\sigma[t]}$. Following the first drug treatment $\sigma[0]$, the treatment-naive matrices $(A^i_{\sigma[t],\text{Naive}})^2_{i=1}$ $(\xi_{\sigma[t]} = 1)$ are used. For the system matrices after the first and any drug switch $\sigma[t_s{-}1] \neq \sigma[t_s]$ thereafter there are two possibilities:

1) if the switch-specific post-treatment dynamics $(A^i_{\sigma[t_s],\text{Post}})^2_{i=1}$ are known, i.e. data for the drug sequence $\sigma[t_s - 1]$-then-$\sigma[t_s]$ is available, then those $(A^i_{\sigma[t_s],\text{Post}})^2_{i=1}$ are used
2) otherwise the best we can do is to sample $\xi_{\sigma[t_s],t_s} \sim F_{\sigma[t_s]}$ and obtain the post-treatment dynamics as $(A^i_{\sigma[t_s],\text{Post}})^2_{i=1} = (A^i_{\sigma[t_s],\text{Naive}} \cdot \xi^i_{\sigma[t_s],t_s})^2_{i=1}$.

Note again that $(A^i_{\sigma[t_s],\text{Post}})^2_{i=1}$ is determined at switching time step t_s and remains fixed until the next drug switch occurs. We make two additional assumptions: a) once the simulated number of living cells gets to or below zero, it stays zero ($x_{T_\text{Death}} \leq 0 \implies x_j = 0 \; \forall j \geq T_\text{Death}$), and b) the shortest possible time between two drug treatments is 12 hours/three time steps ($t_{s'} - t_s \geq 3$).

The simulator was tested on two-drug schedules for which experimental data is available. The results resembled the qualitative and quantitative behavior of the measurements.

B. Drug Schedule Control Approaches

To design a finite time drug schedule multiple competing objectives have to be considered, we consider four: a) increase the efficacy in reducing the number of cancer cells, b) reduce toxicity to non-cancer cells, c) decrease the risk of developing drug resistance, and d) infrequent switching of drugs so that the drug schedule can be administered more easily in practice. One possible approach is to formulate drug scheduling as optimal control problem over the sequence of drugs and switching times. However, the combinatorial nature of the optimization and stochasticity in the dynamics shift (F_δ) are challenging. Hence, for our initial exploration, we introduce a baseline controller and four variants of a heuristic based closed-loop controller.

1) Two-Drug Baseline Controllers: One natural approach is to take the two most cytotoxic (able to kill cancer cells fast) drugs and apply them alternately. Varying the number of drug switches $N_S \in \{1, 2, 8\}$ over a finite time horizon T and thereby the time between drug applications gives us a set of reasonable open-loop baseline controllers (denote BL).

The highly effective drugs δ_1, δ_2 were chosen using the simulator to simulate the number of living cells for every drug in a naive-treatment setting without process noise over 10 time steps (40 hours). We found that the two drugs that reduced the number of living cells the most are the combinations $\delta_1 =$ NLi & Trametinib and $\delta_2 =$ JQ1 & BEZ.

2) Closed-Loop Controller: We explore closed-loop control approaches with two essential building blocks. First, the decision if another drug should be given, and second if so which drug should be given next (see Algorithm 1).

Algorithm 1: Pseudo-code Closed Loop Controller

```
Result: Observed cell counts: x, Drug sequence: σ
initialize cell_experiment(x_init)
for t ← 0 to T do
    x, σ = cell_experiment.get_state()
    if ẋ > 0 and ẍ > 0 and
     σ[t] = σ[t − i] ∀i ∈ {1, 2, 3} then
        next_drug = select_next_drug(x, σ)
        σ[t + 1] = next_drug
    else
        σ[t + 1] = σ[t]
    end
    cell_experiment.apply_drug(σ[t + 1])
end
return x, σ
```

To minimize toxicity and maximize drug switching times, a new drug is only applied after at least 3 time steps after the most recent drug and if accelerating growth is observed ($\dot{x}, \ddot{x} > 0$) or as a less conservative variant, when the population is growing ($\dot{x} > 0$). We estimate $\dot{x}$, and $\ddot{x}$ with finite difference methods. To reduce the risk of developing drug resistance, the new drug has to be different than the previous drug, so we need a drug selection mechanism. Two different drug selection mechanisms $select\_next\_drug(x, \sigma)$ are tried: a) randomly sample from the set of drugs (which

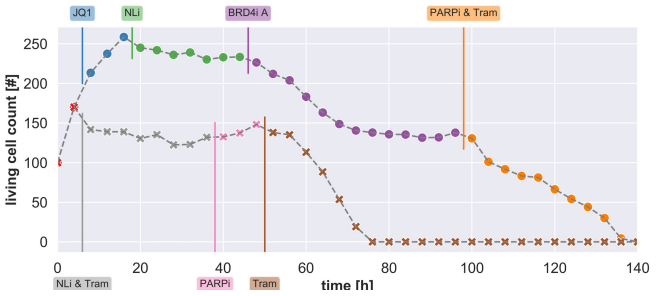

Fig. 5: Two exemplary drug schedules and induced simulated cell count for the $\dot{x}$ controllers where the next drug is randomly selected (top plot) and using $H = 3$ step look-ahead (bottom plot). Vertical lines correspond to the application of the respective drug and the colors after them signify that the cell population evolves under the drugs' influence.

we denote by R), and b) use the drug-specific treatment-naive models to predict the population evolution H steps forward and select the drug that has the lowest predicted cell count after H steps (which we denote by LH for look-ahead). Note that the latter is similar in spirit to the model predictive control approach that [7] designs to mitigate HIV mutation. This yields four distinct controllers: 1) $\dot{x}, \ddot{x} > 0$ R, 2) $\dot{x} > 0$ R, 3) $\dot{x}, \ddot{x} > 0$ LH, and 4) $\dot{x} > 0$ LH.

C. In silico Experiments to identify Promising Schedules

This section presents how we evaluate the set of proposed controllers *in silico*, how identify promising drug schedules *in silico*, and the results of *in vitro* experiments.

1) In silico Controller Evaluation: Fig. 5 shows example trajectories of a cancer cell population for two $\dot{x} > 0$ closed loop controllers: random and $H = 3$ step look-ahead next drug selection for the same initial conditions. The controllers were evaluated initializing the simulator with 24 different realistic conditions, taken from our DMSO data sets for a finite time horizon of 35 time steps ($140h$). For each initial condition, 1000 runs were performed to account for the sources of randomness: in drug selection, the process noise $\eta_{\sigma[t]}$, and the sampled post-treatment dynamics uncertainty $\xi_{\sigma[t]}$. The resulting trajectories x of the number of living cells were evaluated according to five metrics. Three metrics aim to capture efficacy: 1) final cell count x_T, 2) the percentage of trajectories $\%_{\text{Success}}$ in which cell count reached zero $x_T = 0$, 3) the average time the count reached zero T_{death} conditioned on $x_T = 0$. We take the total number of treatments applied N_{Treat} (e.g. drug sequence Tram-NLi-Tram-BEZ: $N_{\text{Treat}} = 4$) as a measure of toxicity, assuming each drug is equally toxic. Lastly, the *maximum number of repetitions* of a drug N_{Rep} (e.g. for Tram-NLi-Tram-BEZ-NLi $N_{\text{Rep}} = 2$) is a proxy for the risk of developing drug resistance as multiple applications of a drug may increase the probability of resistance. Table I contains the mean and standard errors across all runs and metrics.

Our results indicate that the look-ahead drug selection rule increases efficacy in terms of $\%_{\text{Success}}$, while a stricter rule to determine when to switch (considering $\dot{x}$ but not $\ddot{x}$) decreases N_{Treat}, which is proxy for toxicity. While the $N_S = 1$ baseline performs well in terms of T_{death}, N_{Treat}, and N_{Rep}, its efficacy in terms of $\%_{\text{Success}} = 4.4\%$ is by far the worst. The preferred controller depends on the relative weight of the objectives, however, our results suggest the preliminary conclusion that the closed-loop controllers outperform the baselines in terms of efficacy-toxicity balance.

Controller	x_T	$\%_{\text{Success}}$	T_{death}[h]	N_{Treat}	N_{Rep}
$\dot{x}, \ddot{x} > 0$ R	74±.9	45	85±.3	4.0	1.3
$\dot{x} > 0$ R	**61**±.8	47	83±.3	4.5	1.4
$\dot{x}, \ddot{x} > 0$ LH	71±.9	62	73±.2	4.1	1.8
$\dot{x} > 0$ LH	66±.9	**66**	73±.2	4.7	2.1
BL $N_S = 1$	181±.6	4.4	**61**±.7	**2.0**	**1.0**
BL $N_S = 2$	985±29	26	108±.3	3.0	2.0
BL $N_S = 8$	164±6	51	82±.3	9.0	5.0

TABLE I: Empirical evaluation of controllers with mean and standard error over 24000 runs on five metrics: final cell count x_T, percentage of successful treatment $\%_{\text{Success}}$, if successful time until 0 living cells T_{death}, number of treatments used N_{Treat}, and the maximum number of repetitions of a drug N_{Rep}. For the last two columns the standard error is omitted as it was below 0.02 and similar for all controllers.

2) Identification of Promising Schedules: Building on the controller evaluation, we use the less restrictive $\dot{x} > 0$ look-ahead controller to identify promising drug schedules which are then evaluated *in vitro* in Sec. V-D. We obtain a set of five promising schedules through the following three steps. First, we create a diverse set of schedules by running the $\dot{x} > 0$ look-ahead controller on the simulator 4 times for each of 24 different realistic initial conditions for 35 time steps (140h). After removing duplicates this leaves us with 95 schedules. Secondly, we evaluate the robustness of these schedules *in silico* by running the stochastic simulator 100 times for each of the 24 initial conditions (2400 runs per schedule), while tracking the metrics x_T, $\%_{\text{Success}}$, and T_{death}. Lastly, we compare the mean of those metrics across runs to select a set of five schedules with high efficacy prioritizing low x_T and T_{death} and a high $\%_{\text{Success}}$. We only considered schedules with four or fewer drugs and timings that are realistic for 6-day experiments. The kill percentages $\%_{\text{Success}}$ of the selected schedules are in the range of 47-72%.

D. In vitro Evaluation of Identified Schedules

We evaluated the five identified schedules *in vitro* on the breast cancer cell line SUM149PT starting the schedules 20h after plating and using 10 replicate wells per schedule. Besides the regular DMSO control, we also ran a drug control to test the response of the cancer population to the drug combinations used by the schedules (BRD4i B & Tram, BRD4i C & Tram, NLi & Tram, PARPi & Tram). To evaluate the schedules and the simulator we compare the cell count trajectories from the experiments with 100 trajectories predicted by the simulator such as in Fig. 6.

The *in vitro* experiments confirmed the effectiveness of the schedules as predicted by the high $\%_{\text{Success}}$ *in silico*: in all five schedules, the number of living cells decreased significantly to less than 10% of the initial cell count. The identified schedules achieve this by leveraging drug combinations which are highly toxic on their own: from

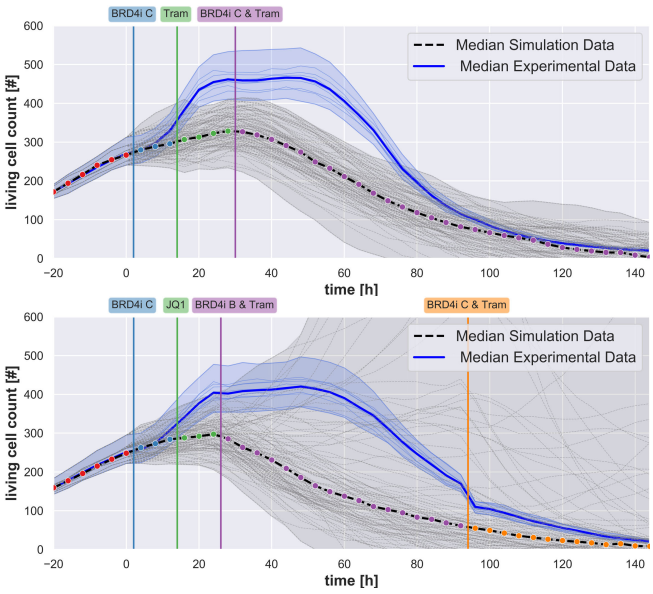

Fig. 6: Experimental and simulation trajectories for two drug schedules identified as promising by *in silico* simulations. The thin lines display the individual trajectories of the 10 experimental wells (blue) and from 100 simulator runs (grey). Thick lines represent the median and the shaded areas the range from min to max values of the trajectories.

the drug controls we find that treating with BRD4i B & Tram, BRD4i C & Tram, or PARPi & Tram over 140h is similarly effective in reducing the cell count close to 10% of the initial. The benefit of the identified schedules is that these toxic drug combinations are applied for shorter times (when a new drug is applied the old one is washed out) which achieves our original goal of improving the efficacy-toxicity balance. At the same time, this also reveals the limitation of our assumption, that all drugs are equally toxic. To further improve the efficacy-toxicity balance we aim to quantify drug toxicity to healthy cells and develop controllers that take toxicity into account e.g. in the form of a constraint.

Comparing the experimental data and the predictions of the simulator as in Fig. 6, we observe that when the sequence of pre- and post-treatment drugs for the schedule were in the initial data, the simulator predictions capture the qualitative behavior well (e.g. top Fig. 6). When the sequence of pre- and post-treatment drug is unknown, hence $\xi_\delta \sim F_\delta$ is sampled, the additional stochasticity leads to highly varying predictions in the simulator (e.g. bottom Fig. 6), which can be interpreted as uncertainty. Note that the median prediction still captures the qualitative behavior measured in experiments. This leads us to two conclusions guiding our future work: Firstly, the extrapolation to unseen sequences of pre- and post-treatment drug via distributions such as F_δ seems promising to capture the range of interaction effects while the median captures the qualitative behavior. Secondly, to improve the quantitative predictions and reduce the uncertainty we need to model a) time-dependence of drug activity, especially if drugs follow shortly after each other (as in Fig. 6), and b) drug-induced permanent changes.

VI. Conclusion and Future Work

In this paper, we presented a general model family informed by biological domain knowledge that accounts for uncertainty induced by pre-treatment. A suitable model was identified using data of two-drug experiments. Further, we presented a set of closed-loop controllers that outperform a two-drug alternating baseline in *in silico* simulation. Lastly, the best closed-loop controller is used to suggest promising schedules which were evaluated in biological experiments. As the next steps, we aim to improve the mathematical model by a) capturing permanent drug-induced changes, and b) time-dependence of drug activity. Furthermore, quantification of toxicity will allow us to design controllers that explicitly take toxicity into account. The code can be found at https://github.com/MariusWiggert/DrugScheduleOpt.

References

[1] D. Barbolosi *et al.*, "Computational oncology-mathematical modelling of drug regimens for precision medicine," 2016.

[2] J. Lehár *et al.*, "Synergistic drug combinations tend to improve therapeutically relevant selectivity," *Nature biotechnology*, vol. 27, no. 7, pp. 659–666, 2009.

[3] B. A. Chabner and T. G. Roberts, "Chemotherapy and the war on cancer," *Nature Reviews Cancer*, vol. 5, no. 1, pp. 65–72, 2005.

[4] M. J. Lee *et al.*, "Sequential application of anticancer drugs enhances cell death by rewiring apoptotic signaling networks," *Cell*, vol. 149, no. 4, pp. 780–794, 2012.

[5] A. Goldman *et al.*, "Temporally sequenced anticancer drugs overcome adaptive resistance by targeting a vulnerable chemotherapy-induced phenotypic transition," *Nature Communications*, 2015.

[6] M. P. Chapman *et al.*, "Modeling differentiation-state transitions linked to therapeutic escape in triple-negative breast cancer," *PLoS Computational Biology*, 2019.

[7] E. A. Hernandez-Vargas *et al.*, "Switching strategies to mitigate HIV mutation," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 4, pp. 1623–1628, 2014.

[8] S. Koplev *et al.*, "Dynamic rearrangement of cell states detected by systematic screening of sequential anticancer treatments," *Cell Reports*, vol. 20, no. 12, pp. 2784–2791, 2017.

[9] M. P. Chapman *et al.*, "On the Analysis of Cyclic Drug Schedules for Cancer Treatment using Switched Dynamical Systems," in *Proceedings of the IEEE Conference on Decision and Control*, 2019.

[10] W. Oduola *et al.*, "Analysis and control of genetic regulatory systems with switched drug inputs," in *3rd IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2016*, 2016.

[11] A. Heydari and S. N. Balakrishnan, "Optimal multi-therapeutic HIV treatment using a global optimal switching scheme," *Applied Mathematics and Computation*, 2013.

[12] B. Zhao *et al.*, "Addressing genetic tumor heterogeneity through computationally predictive combination therapy," *Cancer Discovery*, 2014.

[13] J. X. Zhou *et al.*, "Nonequilibrium population dynamics of phenotype conversion of cancer cells," *PLoS ONE*, 2014.

[14] V. D. Jonsson *et al.*, "Novel computational method for predicting polytherapy switching strategies to overcome tumor heterogeneity and evolution," *Scientific Reports*, 2017.

[15] N. Totis *et al.*, "A population-based approach to study the effects of growth and division rates on the dynamics of cell size statistics," *IEEE Control Systems Letters*, vol. 5, no. 2, pp. 725–730, 2020.

[16] K. Garg and D. Panagou, "Finite-Time Stability of Switched and Hybrid Systems with Unstable Modes," no. November, 2019. [Online]. Available: http://arxiv.org/abs/1901.08513

[17] J. P. Jansch-Porto and G. E. Dullerud, "Decentralized Control of Switched-Systems with Path-Dependent 2-induced Bounds," *Proceedings of the American Control Conference*, pp. 502–507, 2018.

[18] R. N. Upton and D. R. Mould, "Basic concepts in population modeling, simulation, and model-based drug development," *CPT: Pharmacometrics and Systems Pharmacology*, 2014.

[19] K. J. Åström and P. Eykhoff, "System identification-A survey," *Automatica*, 1971.