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. Mathematical bounds on the difference between this new model and a standard switched system model are provided. 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. The controllers outperform a baseline controller by achieving an improved balance between efficacy and toxicity.

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]–[3]. It is standard practice to develop a single drug model to determine an optimal treatment dose. However, applying multiple drugs in a sequence has the potential to effectively stabilize the cancer population while minimizing toxicity [4] and mitigating the risk of developing drug resistance [5], [6].

Our previous work modeled the cancer cell population response to a given sequence of drugs as a switched dynamical system, using data of the response to a single drug to identify the dynamics [7]. One of the key assumptions was that the error due to neglecting pre-treatment is sufficiently small to allow for asymptotic cell population decay without having the data to justify this claim. In the current paper, we improve upon our former approach in particular by using newly available time-series data of a population’s response to sets of two drugs applied sequentially. The data indicates that the population’s response to a given drug can be altered by pre-treatment with a different drug (see Fig. 1).

Fig. 1: This figure shows time-series data from two separate experiments of the response of breast cancer cell populations to two drugs (Trametinib and NLi) applied sequentially (t=0, t=72h). The line represents the empirical mean of 6 replicate wells, and the shaded area the empirical standard deviation. Observe, that the population dynamics after NLi treatment are different when applied at time zero compared to after 72 hours of Trametinib pre-treatment.

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 on response, is an open research area. In this paper, we use time-series data of breast cancer cell populations treated with 9 drugs in various two-drug treatment regimens, where each condition lasts 72 hours. We use this novel data set to quantify how the cancer 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 therapeutic control.

Related Work. Previous work has been done on modeling cell populations as linear or nonlinear systems [4], [6]. In addition, many strategies have been explored to mathematically identify optimal treatment regimens to manage cancer, HIV, and other illnesses [8]–[10] where it is especially important to minimize drug dose. Control strategies that result in more efficient treatment (for example, less total drug volume applied) can be found by solving linear matrix inequalities using convex optimization techniques, although these results can be overly conservative [8], [9]. The type of model used for cell populations or dynamics can also vary widely, including biological networks [9], [11], nonlinear models [2], [10], or systems of ordinary differential equations [12]. 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 to representation as a switched system. There has been extensive work on the stability and control of these systems [13]–[15].

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 provide mathematical bounds for the worst-case finite time difference between our model considering pre-treatment effects and a model (Sec. [III-C]) neglecting pre-treatment. Further, we use our data-driven models of cell population dynamics with a sequence of a selection of up to nine drugs to build a simulator on which drug schedules can be tested in silico (Sec. [V]). Lastly, we present preliminary work on designing closed-loop drug schedule controllers and evaluate their performance in simulation (Sec. [V]). 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. [I]). 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 (cell death). The second measurement is available 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 its rationale. Moreover, we bound the worst-case difference between our model incorporating post-treatment uncertainty and a treatment-naive switched systems model.

A. Model Family

We represent the evolution of a cancer cell population under a drug treatment schedule as an uncertain, discrete-time, time-invariant, switched linear auto-regressive dynamical system, with the state at time $t$ represented by the $n$ dimensional non-negative state vector $x_t \in \mathbb{R}_+^n$. In this paper we consider $n \in \{1,2\}$. 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 $x_t = [\#living, \#dying]^T$, taken from the available measurements (Sec. [II]). The current state $x_t$ depends on the $p$ previous states $(x_{t-1}, x_{t-2}, \ldots, x_{t-p})$. A drug treatment schedule, denoted by $\sigma$, is a mapping from time $t$ to the most recent drug applied at or before time $t$, hence $\forall t \sigma(t) \in D$, 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 length $T$ 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:

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

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 $A_{\sigma(t)}^{k}$ 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 $\eta_{\sigma(t)} \in \mathbb{R}^{n}$. The multiplicative term $\xi_{\sigma(t), t} \in \mathbb{R}$ 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 $\xi_{\sigma(t), t}$ as the post-treatment dynamics uncertainty. Please note that the multiplicative term $\xi_{\sigma(t), t}$ 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 $(A_{\sigma(t)}^{k})_{k=1}^{p}$ encode how states can affect each other. Observe that for $x_t \in \mathbb{R}$ the matrices $(A_{\sigma(t)}^{k})_{k=1}^{p}$ 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) $A_{\sigma(t), 01} = 0$, meaning the number of dying cells does not influence the future number of living cells, and c) $A_{\sigma(t), 01} = 0$ and $A_{\sigma(t), 11} = 0$, 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 three equally-spaced time windows $N_{tw} \in \{1, 2, 3\}$. $N_{tw} = 1$ represents one mode of drug activity; $T_{w} = [0h, 72h] = \{0h, 4h, \ldots, 72h\}$, for $N_{tw} = 2$ we have: $T_{w}^{1} = [0h, 36h]$ and $T_{w}^{2} = [36h, 72h]$ and lastly for $N_{tw} = 3$: $T_{w}^{i} = [(12i)h, (12i+1)h]$ $\forall i \in \{1, 2, 3\}$.

Assumptions: We make the following modeling assumptions regarding $\eta$ and $\xi$. The drug-specific process noise $\eta_{\sigma(t), t}$ is assumed to be bounded and distributed as a truncated normal 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 $A_{\sigma(t)}^{j}$ represent treatment-naive dynamics, consequently $\xi_{\sigma(t), t} = 1$ 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 of the second drug treatment, onward $\xi_{\sigma(t), t} := \xi_{\sigma(t), t}^{1}, \ldots, \xi_{\sigma(t), t}^{p}$ takes some value drawn at the respective switching time $t_{s}$ 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, $A_{\sigma(t), t}^{j} \cdot \xi_{\sigma(t), t}$ represents the post-treatment dynamics of the cancer cell population. Note that $\xi_{\sigma(t), t}$ is drawn only at the respective switching time $t_{s}$. Consequently, $\xi_{\sigma(t), t} = \xi_{\sigma(t), j}$ for all $j$ after $t_{s}$ but prior to the next drug switch. We assume that the effect of pre-treatment is modeled completely by the multiplicative term $\xi_{\sigma(t), t}$. This is an appropriate modeling choice since the matrices are low-dimensional in this paper, and we will investigate other possibilities in future work as additional high-quality data becomes available.

We estimate the treatment-naive and post-treatment matrices $(A_{\sigma(t)}^{k})_{k=1}^{p}$, 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.

B. Modeling Rationale

The choice of the model family is justified as follows:

1. Linear discrete-time auto-regressive models are sufficiently general to represent 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 multiplicative form of $\xi_{\sigma(t), t}$ is mathematically convenient for analysis to bound the difference between our model incorporating post-treatment dynamics uncertainty and a treatment-naive switched system model, which will be provided in Sec. [III-C]

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 $A_{\sigma(t)}^{j}$. 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 [16]. Hence we divide the 72 hour time frame over which a specific drug is active in $N_{tw} \in \{1, 2, 3\}$ time windows.

C. Worst-Case Finite-Time Difference Analysis

We are interested in studying the difference between our model [1] with bounded process noise $\eta_{\sigma(t)}$ and bounded post-treatment dynamics uncertainty $\xi_{\sigma(t)}$ and a treatment-naive switched system model ($\eta_{\sigma(t)} = 0$, $\xi_{\sigma(t), t} = 1$). Here we bound the worst-case accumulation of this difference over a finite time horizon $T$ for a given schedule $\sigma$. We perform the analysis for a model of auto-regressive order $p = 2$ to simplify the presentation of the mathematics and because of our system identification results, which will be presented in Sec. [IV] indicate the suitability of this auto-regressive order. Note that the analysis can be extended readily to a general auto-regressive order $p$.

For compact notation, we define $\bar{A}_{\sigma(j)}$, $\bar{\xi}_{\sigma(j)}$, $\bar{\eta}_{\sigma(j)}$, and their corresponding norms as follows:

$$\bar{A}_{\sigma(j)} := \begin{bmatrix} A_{\sigma(j)}^{1} & A_{\sigma(j)}^{2} \\ I & 0 \end{bmatrix}, \theta_{\sigma(j)} := \|\bar{A}_{\sigma(j)}\|_{\infty}$$

$$\bar{\xi}_{\sigma(j)} := \begin{bmatrix} \xi_{\sigma(j), 1} & \xi_{\sigma(j), 2} \\ 0 & I \end{bmatrix} A_{\sigma(j)}^{1}, \gamma_{\sigma(j)} := \|\bar{\xi}_{\sigma(j)}\|_{\infty}$$

$$\bar{\eta}_{\sigma(j)} := \begin{bmatrix} \eta_{\sigma(j), 1} & \eta_{\sigma(j), 2} \\ 0 & I \end{bmatrix}, H := \max_{j}(\|\bar{\eta}_{\sigma(j)}\|_{\infty}).$$

We choose the infinity norm because this results in $\gamma_{\sigma(j)} \geq 1$ and $\theta_{\sigma(j)} \geq 1$, which is useful for the analysis.
Lemma 1 (Worst-Case Finite-Time Difference Bound):
Let $\sigma$ be given. Define $\Gamma := \prod_{j=0}^{T-1} \gamma_{\sigma(j)}$, $\Theta := \prod_{j=0}^{T-1} \theta_{\sigma(j)}$, and $X_0 := \|\bar{x}_0\|_{\infty}$. For each $j$, assume that $\xi_{\sigma(j)}$ is independent of $\eta_{\sigma(j)}$. Then, an upper bound on the accumulation of modeling differences up to time $T$ is given by:

$$\|\hat{x}_T - \bar{x}_T\|_{\infty} \leq (\Gamma + 1) \Theta X_0 + (\Gamma \Theta (T - 1) + 1) H,$$

Proof: By propagating the system dynamics \(i\) forward to time $T$, one can show the following equation:

$$\|\hat{x}_T - \bar{x}_T\|_{\infty} = \prod_{j=0}^{T-1} \xi_{\sigma(j)} A_{\sigma(j)} \bar{x}_0 + \prod_{i=0}^{T-1} \xi_{\sigma(j)} A_{\sigma(j)} \eta_{\sigma(i)} + \eta_{\sigma(T-1)} - \prod_{j=0}^{T-1} A_{\sigma(j)} \bar{x}_0\|_{\infty},$$

where $\hat{x}_T = [\hat{x}_T\ \hat{x}_{T-1}]^T$ is the solution to \(i\) at time $T$, and $\bar{x}_T = [x_T\ x_{T-1}]^T$ is the solution to \(i\) with $\eta_{\sigma(i)} = 0$ and $\xi_{\sigma(i),t} = 1$ at time $T$. By taking norms and using the triangle inequality, one can show that

$$\|\hat{x}_T - \bar{x}_T\|_{\infty} \leq \prod_{j=0}^{T-1} \gamma_{\sigma(j)} \theta_{\sigma(j)} \|\bar{x}_0\|_{\infty}$$

$$+ \prod_{j=1}^{T-1} \gamma_{\sigma(j)} \theta_{\sigma(j)} \|\hat{x}_0\|_{\infty}$$

$$+ \cdots + \prod_{j=T-1}^{T-1} \gamma_{\sigma(j)} \theta_{\sigma(j)} \|\hat{x}_{T-2}\|_{\infty}$$

$$+ \|\hat{x}_{T-1}\|_{\infty} + \sum_{j=0}^{T-1} \theta_{\sigma(j)} \|\bar{x}_0\|_{\infty}.$$ 

Since $\gamma_{\sigma(j)}, \theta_{\sigma(j)} \geq 1 \forall j$, the following inequalities hold:

$$\prod_{j=T-1}^{T-1} \gamma_{\sigma(j)} \theta_{\sigma(j)} \leq \prod_{j=T-2}^{T-1} \gamma_{\sigma(j)} \theta_{\sigma(j)} \leq \cdots \leq \prod_{j=0}^{T-1} \gamma_{\sigma(j)} \theta_{\sigma(j)}.$$

Finally, substituting $\Gamma, \Theta, H$, and $X_0$, we find

$$\|\hat{x}_T - \bar{x}_T\|_{\infty} \leq \Gamma X_0 + \Gamma \Theta |\hat{x}_0|_{\infty}$$

$$+ \cdots + \Gamma \Theta |\hat{x}_{T-2}|_{\infty} + \|\hat{x}_{T-1}\|_{\infty} + \Theta X_0$$

$$\leq (\Gamma + 1) \Theta X_0 + (\Gamma \Theta (T - 1) + 1) H.$$ 

We also explored a stochastic difference bound since it is less conservative than a worst-case bound. The stochastic analysis leads to a term involving an expectation of products of $\xi_{\sigma(j)}$ over multiple time points $j$, which cannot be readily simplified. The terms $\xi_{\sigma(j)}$ for all $j$ are not necessarily independent, and further assumptions on the distributions $F_{\sigma(j)}$ are not well-justified by the available data. In Sec. IV-C we will use the empirically identified bounds of $\eta$ and $\xi$ to compute the worst-case difference bound provided by Lemma 1 for example drug schedules.

A. System Matrices

1) System Identification Methodology: We identify the system matrices $(A_{\delta,j})_{i=1}^{P}$ 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 $\delta$ there is one treatment-naive condition and multiple post-treatment conditions. We will identify their respective system matrices by $A_{\delta,Naive}^{j}$ and $A_{\delta,Post}^{j}$.

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 (66-80% of all available wells $W_{Total} = W_{Train} \cup W_{Test}$). For example, we use the experimental data summarized in the upper curve of Fig. 1 until $T_{s} = 72$ hours to estimate $(A_{\delta,Naive})_{i=1}^{P}$ for $\delta = \text{Trematinib}$, whereas we use the data in the lower curve after $T_{s} = 72$ hours to estimate $(A_{\delta,Post})_{i=1}^{P}$ for the NLI post-treatment condition. Under our assumption of additive zero-mean truncated Gaussian process noise, the least-squares loss can be derived from the maximum likelihood inference principle [17]. Recall that an element of the model family is fully specified by a tuple of four variables $(N_{\text{well}}, p, n, C)$. For a given drug $\delta \in D$ with data from the set $W_{Train}$, we fit the model for the time window $W_{Train} = [T_{L}, T_{U}]$ by solving the following least-squares regression problem:

$$(A_{\delta})_{i=1}^{P} = \arg\min \sum_{j=1}^{P} \sum_{w \in W_{Train} = [T_{L}+p]} ||x_{t, w} - \sum_{i=1}^{n} A_{\delta}^{j} x_{t-i, w} ||_{2}^{2},$$

where $(A_{\delta})_{i=1}^{P}$ is subject to the set of constraints $C$, $x_{t, w} \in \mathbb{R}_{+}^{n}$ is the measurement at time $t$ 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-C.

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 \(i\) are used to predict the $k$-following counts of living cells. Fig. 2 shows the predictions of the \((N_{\text{well}}=1, p=2, n=1, C=\text{None})\) model for $k = 3$. We want to choose a high $k$ because the simulator and controllers, which will be presented in Sec. IV, 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 $N_{\text{well}} = 3, p = 3$ 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 \(N_{tw} \in \{1, 2, 3\}\), the number of states \(n \in \{1, 2\}\), the auto-regressive order \(p \in \{1, 2, 3\}\), and the constraints \(C \in \{\text{None}, A_{01} = 0, A_{11} = 0 & A_{11} > 0\}\).

The cross-validation results presented here 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 the other experiments (Sec. III). We fit a \((N_{tw}, p, n, C)\)-model to the treatment-naive portion of the experiment (i.e., from 0 hours to 72 hours), and we fit a \((N_{tw}, p, n, C)\)-model to the post-treatment portion of the experiment (i.e., from 72 hours to 144 hours). For example, Fig. 2 shows the performance of the \((N_{tw} = 1, p = 2, n = 1, C = \text{None})\)-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 and not only in aggregate (e.g. for JQ1/BEZ a specific model tuple could 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 list.

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 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 \(N_{tw} = 2\) but increases for \(N_{tw} = \{1, 3\}\). The latter suggests overfitting of the model on the training data.

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 \(C \in \{\text{None}, A_{01} = 0, A_{11} = 0 & A_{11} > 0\}\), the latter two result in the same MSE, hence the \(A_{11} > 0\) constraint is not active and only \(A_{11} = 0\) matters. The constraint supports generalization for \(p \geq 1\), \(N_{tw} \geq 1\) 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 drug population dynamics models (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 \(N_{param} = N_{tw} \cdot p \cdot n^2 - n \cdot I_{constrained}\) where \(I_{constrained} = 1\) if \(C\) contains \(A_{11} = 0\) 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 \(N_{tw} \in \{2, 3\}\) has superior testing performance, we choose \(N_{tw} = 1\) since this choice reduces the complexity of the model for simulation and control (Sec. V).

Hence, our chosen model is \((N_{tw} = 1, p = 2, n = 1, C = \text{None})\), which will be considered fixed hereafter.

\(^{1}\)Representing drug-treated cell population dynamics with models capturing the time-dependence of drug activity is a promising direction, which we will investigate in future work.
B. Estimation and Results for $\Lambda_\delta$ and $F_3$

1) Process Noise Covariance Matrix $\Lambda_\delta$: The $n$ process noise variances $\nu_j$, $j \in \{1, 2\}$ 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

$$r_{w,t} := x_{t,w} - \sum_{i=1}^{p} A_i^\delta \cdot x_{t-i,w}$$

(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 $t \in T_w = [T_i, T_j]$ resemble a distribution, on which for each state $j \in \{1, 2\}$, normal distribution is fitted to obtain the variances $\nu_j$. To obtain the truncation bounds symmetric around $\mu = 0$, we take the $95$ percentile of the set of absolute residual values.

2) Post-treatment Dynamics Uncertainty Distribution $F_3$: The multiple sets of post-treatment dynamics system matrices $(A_{\delta, \text{Post}}^\delta)^P_{i=1}$ of a drug can be estimated using the procedure described in Sec. IV-A. We can think of the matrices $(A_{\delta, \text{Post}}^\delta)^P_{i=1}$ 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-Tametinib, and PARPi-after-PARPi.

Our experimental data for a specific drug $\delta$ only provides measurements for certain post-treatment conditions. Hence, to estimate the distribution $F_3$, we make two assumptions: a) the post-treatment dynamics of $\delta$ for other drugs or application times are comparable to the ones observed in our experiments, and b) the post-treatment dynamics are unknown, so $\xi_{\delta,t}$ is sampled at each drug switch $t_s$ to estimate the dynamics $(A_{\delta, \text{Post}}^\delta)^P_{i=1} = (A_{\delta, \text{Naive}}^\delta, \xi_\delta)^P_{i=1}$. These assumptions are reflected in the estimation procedure for $F_3$, which is done via bootstrapping of experiment wells. The two steps for obtaining $N_{\delta, \text{post}}$ samples of $\xi_\delta$, where $N_{\delta, \text{post}}$ is the number of experiments in which $\delta$ is given as second drug are:

1) Estimate $(A_{\delta, \text{Naive}}^\delta)^P_{i=1}$ and $N_{\delta, \text{post}}$ the multiple different $(A_{\delta, \text{Post}}^\delta)^P_{i=1}$ using bootstrap on the wells of the respective experiments

2) Calculate for each of the $N_{\delta, \text{post}}$ post-treatment matrices $(A_{\delta, \text{Post}}^\delta)^P_{i=1}$ an $\xi_\delta = (\xi_\delta^1, \ldots, \xi_\delta^P)^T$ using $\xi_\delta = \arg \min_{\xi \in \mathbb{R}^{p \times 1}} \| A_{\delta, \text{Naive}}^\delta \cdot \xi_\delta - A_{\delta, \text{Post}}^\delta \|_2^2$ 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 $\xi_\delta$ to which we fit a kernel density estimator to obtain $F_3$. We sample from this kernel density estimator in the simulator, which will be presented in Sec. V.

3) Estimation Results of $\Lambda_\delta$ and $F_3$: We identified the empirical covariance $\Lambda_\delta$ and distribution $F_3$ for each drug $\delta$ using our chosen model of $(N_{\text{data}} = 1, p = 2, n = 1, C = \text{None})$. This was done to validate the truncated Gaussian assumptions on $\eta_\delta$ and get insight into the post-treatment uncertainty distribution $\xi_\delta \sim F_3$. We found that $\eta_\delta$ is truncated Gaussian for all drugs mostly with $\mu \approx 0$ and varying variances. For $\xi_{\delta,t} = (\xi_\delta^1, \xi_\delta^2) \sim F_3$ the observed empirical $F_3$ shows a strong correlation between $\xi_\delta^1$ and $\xi_\delta^2$ 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.

C. Empirical Worst-Case Difference Bound

For every drug $\delta$, we use the identified treatment-naive system matrices $(A_{\text{Naive}, \delta}^\delta)^2_{i=1}$, $\Lambda_\delta$, and the samples used to fit $F_3$ to estimate $\theta_\delta$, $\gamma_\delta$, and $\| \eta_\delta \|_\infty$, which are needed for the bound in Lemma 1. The estimate of $\theta_\delta$ is available from the treatment-naive system matrix. The estimate of $\| \eta_\delta \|_\infty$ is available from the truncation bounds of the truncated normal distribution. To estimate $\gamma_\delta$, we take the samples used to fit $F_3$ and calculate $\| \xi_\delta^1 \|_\infty$ for each sample $j$. Then, $\gamma_\delta$ is the $95$ percentile of the set of $\| \xi_\delta^1 \|_\infty$ over the samples $j$.

Using those values, the worst-case bound was calculated for the two schedules in Fig. 5. For the top schedule, we observe $\| \tilde{\gamma} - \gamma \|_\infty \leq 1.2 \cdot 10^{-10}$ and for the bottom schedule $\| \tilde{\gamma} - \gamma \|_\infty \leq 1.2 \cdot 10^{-30}$. These bounds are extremely high compared to the initial cell count ($X_0 = 160$). While the individual estimates $\gamma_\delta \in [1.0, 4.7]$ and $\theta_\delta \in [1.0, 2.9]$ are small, $\Gamma$ and $\Theta$ grow exponentially over the time horizon $T = 34$ steps. The high magnitudes emphasize the importance of modeling pre-treatment effects and the need to derive a less conservative difference bound in future work.

V. PRELIMINARY EXPLORATION OF DRUG SCHEDULING CONTROLLERS

The long-term aim of this work is to identify promising treatment strategies in silico, thereby effectively guiding laboratory research on future experiments and improving on the current trial-and-error paradigm. Ultimately, we hope to apply this line of work to inform the design of drug schedules for cancer patients. This section presents preliminary 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 lastly, in silico experimental results 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 $\eta(t)$. Following the first drug treatment $\sigma(0)$, the treatment-naive matrices $(A^1_{\sigma(t)},Naive)^2_{i=1}$ $(\xi(t) = 1)$ 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 $(A^1_{\sigma(t_s),Post})^2_{i=1}$ are known, i.e. data for the drug schedule $\sigma(t_s - 1)$-then $\sigma(t_s)$ is available, then they are used ($\xi(t) = 1$)

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 $(A^1_{\sigma(t_s),Post})^2_{i=1} = (A^1_{\sigma(t_s),Naive} \cdot \xi_{\sigma(t_s),t_s})^2_{i=1}$

Note again that $(A^1_{\sigma(t_s),Post})^2_{i=1}$ is determined at switching time $t_s$ 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 ($x_{T_{Death}} \leq 0 \Rightarrow x = 0 \forall t \geq T_{Death}$), and b) the shortest possible time between two drug treatments is 12 hours/three time steps ($t_{s} - t_{s-1} \geq 12h$).

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

**B. Drug Schedule Control Approaches**

To design a good finite time drug schedule for cancer treatment 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 as this corresponds to less effort for doctors and fewer hospital visits for patients. In an attempt to formulate this problem, we introduce an open-loop baseline controller and four variants of a closed-loop feedback 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 $\delta_1, \delta_2$ 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) _Open-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]).

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 ($\dot{x} > 0$). To reduce the risk of developing drug resistance, the new drug has to be different than the previous drug and a drug selection mechanism is applied. Two different drug selection mechanisms $select_{next\_drug}(x, \sigma)$ are tried: a) randomly sample from the set of drugs (which 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 [6] designs to mitigate HIV mutation. This yields four distinct controllers: 1) $(\dot{x}, \ddot{x} > 0)$ R, 2) $(\dot{x}, \ddot{x}) > 0$ R, 3) $(\dot{x}, \ddot{x} > 0)$ LH, and 4) $(\dot{x}, \ddot{x}) > 0$ LH. The estimation of $\dot{x}$, and $\ddot{x}$ is done using finite difference methods.

**C. Experimental Results**

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 over 24 different realistic initial conditions, taken from our DMSO data sets for a finite time horizon of 35 time steps ($T = 140h$). For each initial condition, 1000 runs were performed to account for the sources of randomness: in drug selection, the process noise $n(t)$, 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 $\%_{success}$ in which cell count reached zero $x_T = 0$, 3) the average time until the count reached zero conditioned on $x_T = 0$. We take the total number of treatments applied $N_{treat}$ (e.g. drug sequence Tram-NLI-Tram-BEZ: $N_{treat} = 4$) 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 $N_{rep} = 2$) is a proxy for the risk of developing drug resistance as multiple applications of a drug may increase the probability of resistance. Table [1] contains the mean and standard errors across all runs and metrics.
The look-ahead drug selection rule, which is proxy for toxicity, while the Naive model was derived to quantify the importance of pre-treatment. A suitable model was identified using data of two-drug experiments. An upper bound for the difference between this model and a treatment-naive model was derived to quantify the temporal differences in drug activity e.g. via a specific parameter or superposition with time-varying multipliers for each drug dynamic. Measurements of the response of healthy cells to drugs would allow for better quantification of toxicity. Lastly, a form of risk-sensitive control such as conditional value-at-risk could be used for drug schedules that account for the uncertainty in post-treatment dynamics. The code can be found at https://github.com/MariusWiggert/DrugScheduleOpt.

**Fig. 5:** Two exemplary drug schedules and induced simulated cancer cell count for the $\dot{x}$ controllers where the next drug is randomly selected (see top plot) and using $H = 3$ step look-ahead (see 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.

<table>
<thead>
<tr>
<th>Controller</th>
<th>$x_T$</th>
<th>$%\text{Success}$</th>
<th>$T_{\text{Death}}$</th>
<th>$N_{\text{Treat}}$</th>
<th>$N_{\text{Rep}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\dot{x}, \dot{x} &gt; 0 \text{ R}$</td>
<td>74 ± 9</td>
<td>45</td>
<td>85 ± 3</td>
<td>4.0</td>
<td>1.3</td>
</tr>
<tr>
<td>$\dot{x}, \dot{x} &gt; 0 \text{ LH}$</td>
<td>71 ± 9</td>
<td>62</td>
<td>73 ± 2</td>
<td>4.1</td>
<td>1.8</td>
</tr>
<tr>
<td>$\dot{x} &gt; 0 \text{ LH}$</td>
<td>66 ± 0</td>
<td>66</td>
<td>73 ± 2</td>
<td>4.7</td>
<td>2.1</td>
</tr>
<tr>
<td>BL $N_S = 1$</td>
<td>181 ± 6</td>
<td>4.4</td>
<td>61 ± 7</td>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td>BL $N_S = 2$</td>
<td>985 ± 29</td>
<td>26</td>
<td>108 ± 3</td>
<td>3.0</td>
<td>2.0</td>
</tr>
<tr>
<td>BL $N_S = 8$</td>
<td>164 ± 6</td>
<td>51</td>
<td>82 ± 3</td>
<td>9.0</td>
<td>5.0</td>
</tr>
</tbody>
</table>

**TABLE I:** Empirical evaluation of controllers with mean and standard error over a finite time horizon $T = 140h$ on five metrics: final cell count $x_T$, percentage of successful treatment $\%\text{Success}$, if successful time until 0 living cells $T_{\text{Death}}$, number of treatments used $N_{\text{Treat}}$, and the maximum number of repetitions of a drug $N_{\text{Rep}}$. For the last two columns the standard error is omitted as it was below 0.02 and similar for all controllers.

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 $\dot{x}$) decreases $N_{\text{Treat}}$, which is proxy for toxicity. While the $N_S = 1$ baseline performs well in terms $T_{\text{Death}}, N_{\text{Treat}}$, and $N_{\text{Rep}}$, its efficacy in terms of $\%\text{Success} = 4.4\%$ is by far the worst. While the preferred controller depends on the relative weight of the objectives, we draw the preliminary conclusion that the closed-loop controllers outperform the baselines in efficacy-toxicity balance. Especially, the less restrictive $\dot{x} > 0$ look-ahead controller is promising.

**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. An upper bound for the difference between this model and a treatment-naive model was derived to quantify the importance of pre-treatment. Lastly, we presented a set of closed-loop controllers that outperform a two-drug alternating baseline on a simulator that we developed.

As a next step, we aim to evaluate the performance of our control approaches in biological experiments. Further, the mathematical model could be improved to better capture the temporal differences in drug activity e.g. via a specific parameter or superposition with time-varying multipliers for each drug dynamic. Measurements of the response of healthy cells to drugs would allow for better quantification of toxicity. Lastly, a form of risk-sensitive control such as conditional value-at-risk could be used for drug schedules that account for the uncertainty in post-treatment dynamics. The code can be found at https://github.com/MariusWiggert/DrugScheduleOpt.

**REFERENCES**


